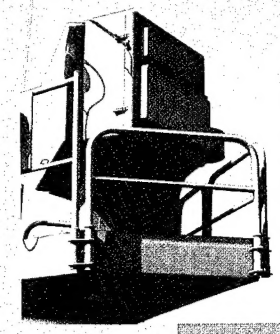
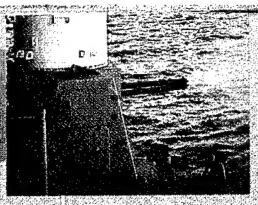
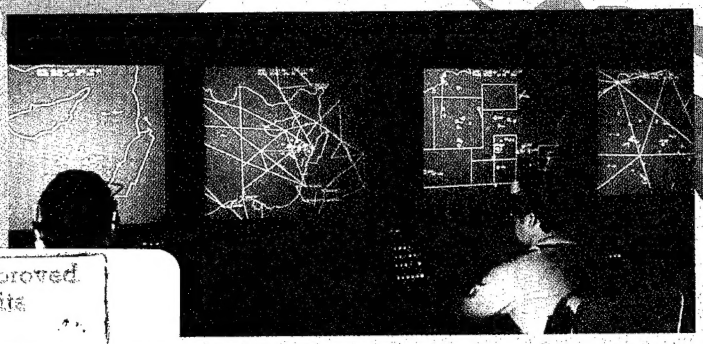
Naval Surface Warfare Center
Dahlgren Division

Technical Digest

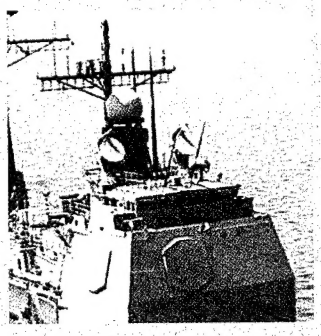
September 1994







Ship Defense Technology



The Naval Surface Warfare Center Dahlgren Division Technical Digest presents unclassified articles, contributed primarily by Division scientists and engineers, on selected research and development programs. The Dahlgren Division, under the leadership of the Naval Surface Warfare Center, provides research, development, test and evaluation, engineering, and fleet support for surface warfare systems, surface ship combat systems, ordnance, mines, amphibious warfare systems, mine countermeasures, special warfare systems, and strategic systems. Please address any correspondence concerning the *NSWCDD Technical Digest* to: Dahlgren Division, Naval Surface Warfare Center, Technical Digest (Code E282), 17320 Dahlgren Road, Dahlgren, VA 22448-5100. Telephone (703) 663-8921.

About the cover: Operation of naval forces in littoral regions of conflict around the globe confronts ships with a new set of stressing threat and environmental requirements. Systems engineering technology is required to meet the operational challenge with affordable solutions. Sensors, decision makers, and weapons must be effectively integrated to allow naval forces the freedom of action required to execute their missions. The Rapid Antiship Missile Integrated Defense System (RAIDS) display shown at the base of the shaded bar area) represents the interface between the integrated system and the decision maker. Continuing clockwise from RAIDS, the SLQ-32 and CIWS represent the engagement system, and each also provides sensors that are integrated into the RAIDS display. The human and machine interface for situational awareness is next illustrated by the CDS display for CG-47 Ticonderoga class cruisers shown in the top right photo. Finally, the SPY-1 radar and the Mk-99 fire-control system are visible in the bottom right photo.

NSWCDD/MP-94/94
Approved for public release;
distribution is unlimited.

Naval Surface Warfare Center
Dahlgren Division

Technical Digest

September 1994

Guest Editor's Introduction:

Towards a Proactive Surface Force—The Role of Ship Defense in the 21st Century

Thomas H. McCants, Jr.

The surface Navy is undergoing a dramatic change of course. This course change is being stimulated primarily by the impact of the end of the Cold War on the missions the Navy must execute and, to a lesser extent, by the influence of advances in antiship technology on the freedom of action required to accomplish those missions. Advancing ship defense technology must supply the means to successfully accomplish that course change if the surface Navy is to remain viable.

The Global Threat Focus

The previous four decades have seen the Fleet focus on global confrontation with a single power possessed with enormous quantitative (if not qualitative) advantages in many theaters of joint interest. The missions considered to stress ship defense requirements were largely open-ocean operations such as European resupply across the Atlantic and major strikes from the ocean through the Greenland-Iceland-United Kingdom gap. The environmental conditions were severe, but little background traffic existed in these places to complicate the engagement decision. The surface forces considered were large, and aircraft carriers were available in each mission to provide air cover and maximize the timeline available for action. Layered defense ensured the readiness of inner defenses, should an encounter with the forces of the opponent prove to be an attack.

The current surface Navy can no longer be characterized as having a single focus or mission. Each operating theater has its own unique blend of environmental conditions, potential opponents, and antiship threat technologies. Though the quantities and qualities of many of these opponents are substantially reduced relative to those of the Soviet Union, the conditions of exposure in many ways make them a greater threat to the ability of the surface Navy to execute its missions.

While those of us whose entire experience stemmed from the Cold War realities might consider the exploding numbers of potential adversaries and operating environments as something new, it may be better seen as a return to traditional Navy roles. The Fleet's mission in the new world order shares many similarities with its pre-World War II mission as the surface Navy is frequently called on to carry out missions widely distributed around the world. The open oceans are used as a passage to surface Navy missions areas, with little likelihood of combat far from land. The surface Navy is frequently the only U.S. force in the area for some period prior to the arrival of joint forces. The surface force on site seldom numbers more than a few ships and frequently

must operate without air cover or even airborne surveillance. The surface Navy must deal with a cluttered near-land environment, where large numbers of friends, neutrals, and hostiles alike pass through the surveillance volume of a ship. The surface Navy must deal with tremendous uncertainty as to who might be a potential belligerent and when hostilities might break out. The nation has demonstrated little tolerance for casualties among U.S. forces or for casualties inflicted upon others due to error.

The result of the Cold War focus was the development and deployment of ship defense systems suited to the expected missions and environments, subject to the limits of the technologies of the day and the realities of integration into the surface Navy infrastructure that evolved after World War II. The defense systems focused on antiair warfare (AAW) and antisubmarine warfare (ASW) in the open ocean. Mine warfare was not a major concern in the open ocean. Surface warfare (SUW) was expected to take place over long distances, with carrier aircraft being the primary instrument. Strike warfare was again the province of the carrier aircraft, later supplemented by cruise missiles. The close-in weapons of the Fleet were expected to be of little use, as the open-ocean war would likely occur over hundreds of miles. Naval guns were essentially abandoned, and short-range missiles were adopted from air-launch weapons (RIM-7 developed from AIM-7) as an inexpensive alternative to developing a specialized self-defense missile for surface ships.

The New Environment—Near-Land Operations

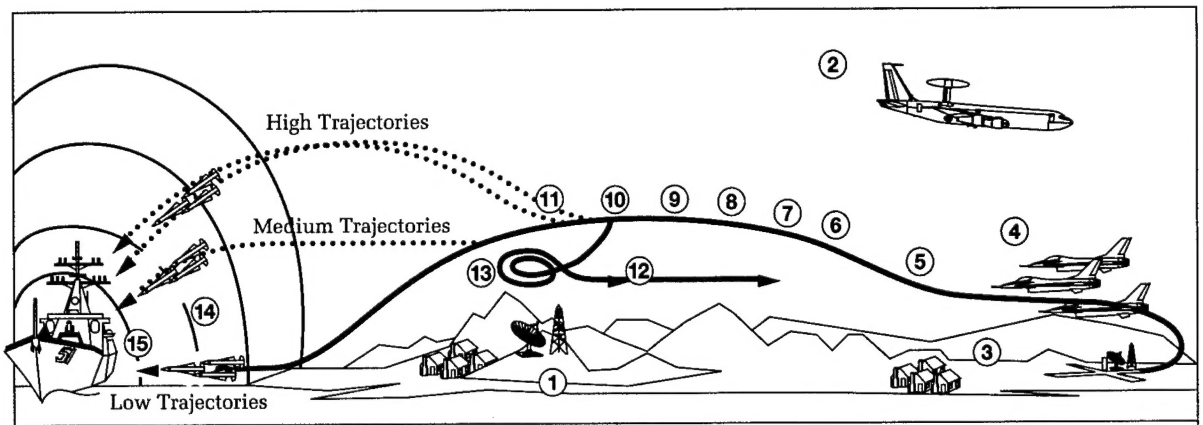
In the new environment, it is frequently difficult to keep potential adversaries at a distance. In fact, the surface Navy must often sail through waters frequented by the adversary. The Earnest Will operations found U.S. warships surrounded by Iranian boats as well as Iranian warships and aircraft. Iranian mines severely damaged *USS Roberts* (a frigate). Surface gun engagements of small Iranian boats were a factor in the *USS Vincennes'* decision to shoot down an Iranian Airbus. Iraqi Exocet missiles severely damaged *USS Stark* despite detection by the *Stark* of the launch aircraft. While STANDARD missiles were fired against an Iranian warship and against Iranian fighter aircraft, no defense was applied against the Iraqi attack. These observations reinforce the argument that short-range, quick-reaction defenses are required for AAW, SUW, and mine warfare. Further, while submarines have not yet been a factor, proliferation of diesel submarines forewarns of the need for short-range, quick-reaction ASW defense as well.

The key requirements of ship defense are: (1) it must effectively defeat attacks, (2) it must do no harm to others, and (3) it must be affordable. The crew must have confidence in their ability to defend themselves against attack if they are to exercise the ship's freedom of action to execute its missions. In order to place their trust in the defense system, and to thus enable the quick reaction required for effectiveness, the crew must have confidence in their ability to control the actions of the defense system. In order for the nation to invest in adequate defense for the ship, the defense must be affordable in light of the value added of the ship and its missions to the defense of the nation.

Defense Against Antiship Capable Missile (ASCM) Threats

To effectively defeat attacks (such as the fighter-launched ASCMs depicted in Figure 1), current AAW defense systems depend on the five cornerstones laid down by RADM Wayne E. Meyer in his challenge to the community responsible for the development of AEGIS: firm track range, reaction time, firepower, coverage, and resistance to degradation. The resulting AEGIS combat system has the capability to deal effectively with the threat that stimulated its development: open-ocean confrontation where the enemy was easy to identify once detected, and where the automatic engagement required for success could be utilized with little concern for inadvertent danger to noncombatants or friends. Self-defense systems developed during this period (Close-In Weapon System (CIWS), NATO Sea Sparrow Missile System (NSSMS), SLQ-32, etc.) had similar foundations requiring automated response for success against the cruise-missile threat in the open ocean.

While the basic cornerstones laid down by RADM Meyer still hold meaning to the designers of the next generation of ship defenses, their influence must be translated for application to a new set of environments and threats. An integrated defense approach involving electronic warfare (EW) interventions, hard-kill (HK) interventions, and HK/EW interactions selected based on attack phase (as depicted in Figure 1 and Table 1) is needed to effectively counter the ASCM near-land threat. Further, the need to ensure that the system does no harm to others and



1. *Surveillance/ Recon Provides Ship Position*
2. *Recon Provides ID, Updates Position*
3. *C³I Engages, Vectors A/C to Acquisition*
4. *Aircraft Fly Low Altitude Run-in*
5. *A/C Climb to Launch Altitude*
6. *Lead Aircraft Starts Search*
7. *A/C Detects Ship and Background*
8. *A/C Selects Target Ship From Background*
9. *Lead A/C Designates Target to Others*
10. *A/C Initialize, Launch First Missiles*
11. *A/C Initialize, Launch Second Missiles*
12. *A/C Dive and Turn Outbound*
13. *Jamming A/C Provide Support for ASM*
14. *ASM's Turn On Seekers*
15. *Seekers Acquire Ship, Complete Attack*

Defense Timeline Can Start With:

External Raid Notification
Skin Track on Aircraft Formation
Aircraft Fire-Control Radar
Aircraft Jamming (if Used)
Skin Track of Missile Separation
Missile Skin Detection
Missile Emissions Detection

Figure 1. *Fighter-launched missile scenario (launch over land).*

Table 1. *EW Interventions, HK Interventions, and HK/EW Interactions*

ATTACK PHASES	INTERVENTIONS (EW)	INTERVENTIONS (HK)	INTERACTIONS (EW TO HK)	INTERACTIONS (HK TO EW)
Surveillance	Counter Surveillance	Attack Surveill Assets	Provides Warning, Prevents Accurate Detect, Provides Time for HK	Forces Surveill From Distance, Improves EW Effectiveness
Reconnaissance	Cover and Deception	Attack Recon Assets	Provides Warning, Prevents Accurate Recon, Provides Time for HK	Forces Recon From Distance, Improves EW Effectiveness
Attack Vectoring	Counter C^3I Comms	Attack C^3I Comms Assets	Provides I&W, Ensures Reaction by HK, Provides Time for HK	Forces Recon From Distance, Improves EW Effectiveness
Acquisition	Counter Acquisition	Kill Platform	Detects Attack, Delays Acquisition, Provides Time for HK	Forces Acq From Distance, Reduces Time for Acq, Improves EW Perf
Targeting	Counter Targeting	Kill Platform	Dilutes Attack, Reduces Targeting Accuracy, Provides Time	Forces Tgt From Distance, Reduces Tgt Accuracy, Improves EW Perf
Launch	Counter Launch	Kill Platform	Provides Time for HK	Forces Launch From Distance, Reduces Aim Accuracy, Improves EW Perf
Weapon Direction and Guidance	Counter Weapon Comms	Kill Platform or Weapon	Forces Platform Closure, Time for HK, Easier HK Trajectory	Forces Platform Standoff, Improves EW Eff, Easier EW vs Guidance Modes
Weapon Target Acquisition	Distraction	Kill Weapon	Dilutes Attack if Kill Recognized, Easier HK Trajectory for EW	Provides Detection for Short Range Seeker Turn-on, Provides Time
Terminal Homing	Seduction	Kill Weapon	Dilutes Attack if Kill Recognized, Easier HK Trajectory	Kill Weapon Not Vulnerable to EW
Fuzing	Counter Fuze	Kill Weapon	Reduces Ship Damage	Kill Weapon Not Vulnerable to EW
Target Damage	Retain Capability	Counter Weapon Effects	Retains Defense Capability Despite HK Damage	Retains Defense Capability Despite EW Damage

the need to ensure that new systems can be afforded expand the cornerstones.

In the new environment, the highly automated defense systems developed for the open-ocean environment have several major disadvantages that new advances must overcome:

- impact of the cluttered near-land environment on the ability of sensors to detect threats
- propensity of sensors in these cluttered environments to promote environmentally derived tracks as if they were actual threats
- crews' awareness of the intolerance for erroneous engagements leading to casualties among friends or others
- intolerance for erroneous or inadvertent firings and resultant casualties
- mismatch between the availability of data upon which to make life-or-death decisions and the time available to make those decisions if an attack is actually in progress
- mismatch among current Fleet weapons and portions of the vastly expanded threat categories to which the fleet may be exposed in near-land environments
- readiness of existing systems limited by aging hardware, computer programs, and support technology
- architecture of these existing systems, making them difficult to support and expensive to modernize.

Change and the rate of change continue to accelerate. We know with certainty that the best system we can design today will be long obsolete before the service life of a ship (25 to 50 years) is ended. We must provide system architectures that facilitate, rather than impede, change.

With the computer upon which I write this paper, adding a new piece of system hardware is so incredibly simple that even I can do it. Adding a scanner requires little more than plugging standard cables in almost any sequence and clicking on an icon on a floppy disk to load specialized software. The computer's system software has no difficulty supporting the new peripheral. The installation documentation required referring to one or two pages of one-sentence descriptions and cartoons. Operation of the scanner is so intuitive and supported by on-line help so well that, after one year, I've yet to resort to the manual.

While this analogy with personal computers may appear simplistic, it is apparent that

future shipboard systems should capitalize on commerical off-the-shelf (COTS) computers for greater modularity and lower cost. The architecture of the system is the key to maintainability, flexibility, and growth. Contrast this to inflexible point-to-point architecture and ancient computer technology aboard many of today's ships and see the source of a new vision for the future.

Threat to Ships From Land Warfare Weapons

The substantial investment made by the Soviets, our allies, and even by ourselves in antiship weapon technology is proliferating around the globe. Further, near-land operations can result in exposure of ships to weapons designed for use in traditional land warfare such as anti-armor attack helicopters and battlefield artillery, including multiple-launch rocket systems and guided projectiles. Short-range and multiwarfare weapons are receiving new emphasis as threat categories expand and threat proximity stresses existing systems. Coordinated use of multiple, layered HK and EW assets is required to achieve high survival probability. Each layer must compensate for the shortcomings of the other layers in order to provide affordable defense (i.e., the silver bullet we seek is actually made of "unobtainium").

Role of Offensive Technology in Ship Defense

Technology has a lot to offer in ship defense. It has even more to offer in establishing new missions for the surface Navy to keep the force pertinent in the next century. For example, threat air defense from the surface force allows long endurance capability to be deployed over wide areas of the globe. Continuous naval surface presence in many operating areas ensures that the capability is likely on-site even if the need emerges quickly. This unobtrusive sea-based capability can be established without requiring a U.S. presence ashore where it might be subjected to local internal political pressures and potential terrorist attack. Expanded exploitation of long-range surface-to-air missile capability (such as modified AEGIS/STANDARD Missile Block IV) could also be used in an offensive manner, such as enforcing "no-flight" zones over the territory of a belligerent, without exposing U.S. airmen to ground-based defense and risk of capture.

Exploitation of global positioning satellite technology and vertical launchers could yield a high-capacity, medium-range strike capability, enabling many target categories to be engaged from the sea at costs lower than those associated with cruise missiles and, again, not forcing exposure of U.S. and allied airmen to ground-based defenses. Advanced gun technology would allow a further multiplication of target capacity as defenses are rolled back, and the ships move closer to shore. The quick-striking missiles and the smaller gun-delivered ordnance would both supplement and enable the particular strengths of cruise missiles and manned aircraft attacks to be effectively exploited. Further, in the absence of aircraft early on in a confrontation, each individual ship would have sufficient attack capability to make U.S. pressure more credible.

A 5-inch gun attack that requires dozens of rounds to be expended to defeat a target means that very few targets can be engaged prior to reloading the ship. Further, if that attack can only occur from 5 to 10 miles, few targets are at risk, and the ship itself is at extraordinary risk. With strike missiles to roll back fixed shore defenses or to strike a few dozen high-value targets, each ship can have an influence. If the gun can fire several tens of miles and reliably hit fixed targets or target areas, hundreds of targets vulnerable to small weapons also become at risk. Thus, a few surface ships properly supported with targeting data sources and located at key trouble spots could exert considerable influence over forces ashore, assuming that the ships themselves are not vulnerable to easy attack from the shore.

With the expansion of offensive capabilities described above, the surface Navy could again assume the status as instrument of national policy that naval forces represented during the centuries preceding the advent of air power. The sitting-duck naval forces subjected to attack by Col. Billy Mitchell in the 1930s no longer exist. The current challenge is not only to ensure that the *USS Stark* and *USS Vincennes* incidents are not repeated, but to ensure that the battle initiative returns to the surface ship. If we meet that challenge, the surface Navy will endure for future generations.

The Technical Digest Articles

This edition of NSWCDD's *Technical Digest* addresses a number of the offensive and defensive facets of future surface combatants discussed in this editorial. The

summary below follows the detect/control/engage functional allocation paradigm again codified by RADM Meyer, modified by the addition of system engineering functions required to support the shipboard functions.

Systems Engineering

The proliferation of fast, highly maneuverable, low-flying antiship missiles will impose a serious threat to U.S. surface ships. Phillips and Chisholm in their article on short-range AAW discuss the threat evolution and its impact on defensive missile systems. They discuss missile guidance and terminal homing options to counter both single and multiple antiship missile attacks. Their analysis shows that decreasing detection range will require fast, highly maneuverable defensive weapons with quick reaction time to allow safe intercept range.

Ervin, in his article, discusses how integrated Interior Communication and Control $(IC)^2$ can improve defensive system effectiveness through improved probability of correct decision-making and reduced reaction time. He describes the ship architecture vision for the future, with emphasis on affordability through the use of commercial off-the-shelf (COTS) equipment. The $(IC)^2$ approach is influenced by (and will influence) the self-defense system architecture. For example, the Rapid Antiship Missile Integrated Defense System (RAIDS) was part of a successful proof-of-concept demonstration of the $(IC)^2$ approach conducted at the NSWC Wallops Island Facility.

Detect

The problem of detecting the low-flying cruise missile threat was the subject of a previous issue of NSWCDD's *Technical Digest* in September 1992. This issue treats other ship defense sensing issues instead. Drescher and Sitzman discuss new methods for tracking TBMs using surface ships. It is noted that similar problems apply to non-TBM high-angle threats and cooperative engagement capability. High-performance surveillance systems are not likely to be affordable for high-elevation coverage for all ships, thus a defense system must rely on remote sensors, such as satellites, for cuing of sensors on the launch ship.

At the other extreme, mines will continue to pose a serious threat to our surface ships. Clem, in his article on superconducting magnetic sensors, responds to the emerging mine threat in near-land missions, particularly buried mines. Despite concern over the missile threat, combat losses of U.S. ships from mines exceeded losses to missiles in the last 10 years. The mine-hunting effectiveness of sensors is stressed by the near-land environment. The article describes the application of advancing technology to improve sensor performance.

Control

A major function of the self-defense control system is to coordinate between weapons and sensors to eliminate interference or at least maximize performance in the presence of interference. Current electromagnetic interference (EMI) management techniques must be applied in advance. Rule-based doctrinal EMI control methods are expected to be provided for the ship self-defense system prevention of conflict between sensors and weapons of the single ship. However, this method will be difficult to extend to the force. Neel, in her article on managing the dynamics of the electromagnetic environment, discusses technology and techniques to be applied in real time across the force.

Engage

Current surface ship systems have layered defense systems including STANDARD Missile for high-altitude long-range engagements; and NSSMS, rolling airframe missile (RAM), and CIWS for close-in defense. Kimbrell, in his article on EW, describes the importance of layered defense, in particular the coordination of HK and EW assets, expanded sensor use of the electromagnetic spectrum, sensor integration and fusion, and EW as weapon of least regret. The use of EW sensors as an aid in target identification, threat, and hostile intent assessment are also discussed.

Simulation as a means of improving training of Navy personnel is an important DoD thrust. Welle et al. describe multiple pulse emulator simulation technology to support countertargeting of threat missiles and countersurveillance of long-range "weapons." In training, the simulator ensures that operators have familiarity with the near-land emitter environment they may see in combat. The simulator development description illustrates the use of COTS.

Gun systems, such as the Phalanx CIWS, are expected to continue to play a major role in ship defense against cruise missiles. Borland discusses the need to expand the capability of Phalanx beyond countering the AAW threat to include the SUW threat posed

by high-speed surface craft. The multiwarfare capability of this system will require the priority between surface and air threats to be managed (either by CIWS or by a surface-ship defense system). Note that many weapons useful for AAW also have SUW capability and, thus, a single warfare area control system may not be cost effective for self-defense.

Short-range multiwarfare weapon effectiveness and the expansion of ship missions drive the need for improvements in gun accuracy. Boyer et al. discuss a nonlinear least-squares method for improved gun accuracy that has been implemented aboard the Arleigh Burke class of AEGIS destroyers. This method can be applied in real time and should be applicable to missile command guidance and to guided projectiles.

Maneuvering threats and much higher speeds, coupled with vertical launch, will stress the ability of missile controls for self-defense. Bibel et al. address robust flight control techniques and their application in designing flight-control systems for surface-launched tactical missiles. Applying these techniques to the design of missile autopilots greatly improves the stability and performance robustness characteristics of the flight-control systems which, in turn, increases the effectiveness of the overall missile system.

While the missiles and guns are expected to provide a formidable defense against the low-flying cruise missile, situations may arise where engagements occur at close range, thereby exposing the ship to fragments or large threat missile components. Higdon describes a water barrier self-defense system against this debris based on water plumes generated by explosive charges simultaneously detonated at a safe standoff from the ship. Analyses and tests demonstrate the viability of this approach in stopping fragments and simulated missile parts.

The Guest Editor

THOMAS H. McCANTS, JR., Principal System Engineer in NSWCDD's Ship Defense Department, earned a B.S. in aerospace engineering from Virginia Polytechnic Institute and State University (VPI&SU) in 1968. Since arriving at NSWCDD in 1964 as a Cooperative Education student trainee, he has performed research and development in missile weapon systems effectiveness and ordnance lethality. He has extensive experience in air target vulnerability, Phalanx system lethality, and TOMAHAWK survivability. From 1977 to 1978, he served as the NSAP Science Advisor to COMOPTEVFOR, Norfolk. He was subsequently the NSWC Program Manager for STANDARD Missile, led the AEGIS Systems Analysis and T&E organization at NSWCDD, was Phalanx CIWS Lead Laboratory Program Manager, and was the Air Target Vulnerability Program Manager at the Division. In 1988 he received the Bernard Smith Award and in 1992-93 received an NSWCDD Fellowship for Postgraduate Studies at VPI&SU. In addition to pursuing an M.S. in Systems Engineering, he is currently Deputy Director of the Navy Theater Ballistic Missile Defense Cost and Operational Effectiveness Analysis. He also supports the PEO (TAD) in systems engineering and analysis for Ship Self-Defense and the Office of Naval Research in a 6.2 Ship Self-Defense program.

Short-Range Antiair Warfare Missile Systems Engineering

Craig A. Phillips and James M. Chisholm

Political and technological developments of the past few years will have a profound effect on future combat situations for the U.S. Navy. Most of the envisioned combat situations will involve unpredictable near-land engagement scenarios. These types of scenarios place an increased emphasis on the short-range antiair warfare (AAW), or self-defense, capabilities of U.S. Navy surface combatants. Ship self-defense capabilities have already been stressed by the development of small, fast, highly maneuverable, low-flying antiship missiles. It is with these factors in mind that the Naval Surface Warfare Center Dahlgren Division (NSWCDD) has participated in missile weapon systems engineering analysis. This article presents some of the analysis that has been performed in the short-range AAW subject area. The dominant technical challenges that encompass the ship self-defense problem are discussed and characterized.

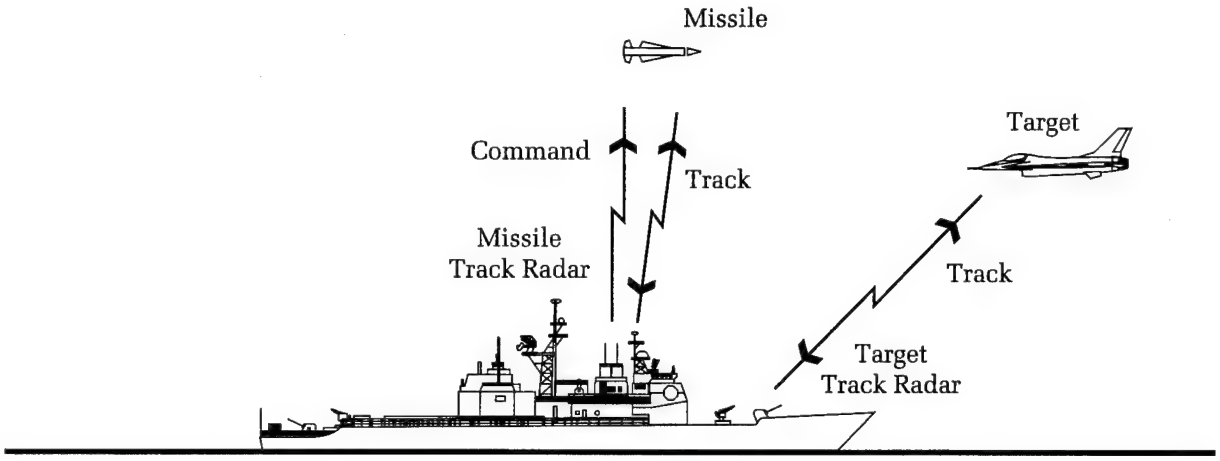
Introduction

The breakup of the Soviet Union and the dissolution of the Warsaw Pact has resulted in a restructuring of priorities within the U.S. Navy. Instead of focusing on a major 'blue-water' confrontation with the Soviet Navy, the emphasis in strategic planning has shifted to littoral, or coastal, warfare scenarios. A maritime strategy of protecting vital U.S. national interests throughout the post-Cold War world, along with budgetary constraints, dictates a need for increased flexibility in applying naval power. This may subject U.S. Navy surface ships to unpredictable engagement scenarios and battle conditions in the future.

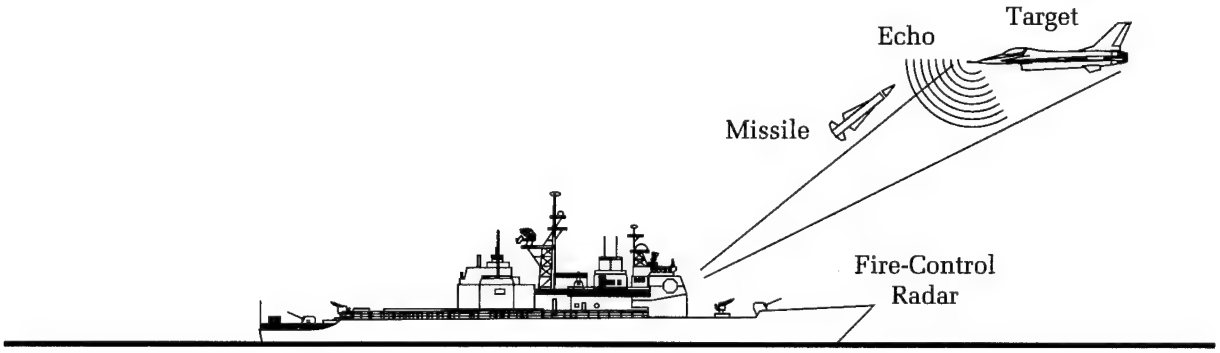
Advances in technology have increased the capabilities of surface-to-surface and air-to-surface antiship missiles. The proliferation of these types of weapons throughout the world, especially to potential adversaries, imposes a threat to all U.S. Navy surface combatants. The British Royal Navy's experience in the 1982 Falklands War and the *USS Stark* incident of 1987 highlight the vulnerability of surface combatants to sea-skimming antiship missiles.

To counter present and future airborne threats to surface combatants, a variety of options are available for a weapon system designer (see Missile Guidance and Control Term Definitions at the end of this article). In this article, three design options for a short-range AAW missile weapon system are considered: (1) a command-all-the-way guidance system, (2) a semiactive home-all-the-way guidance system, and (3) a combined command midcourse/terminal homing guidance system. The options, illustrated in Figure 1, will now be discussed in detail.

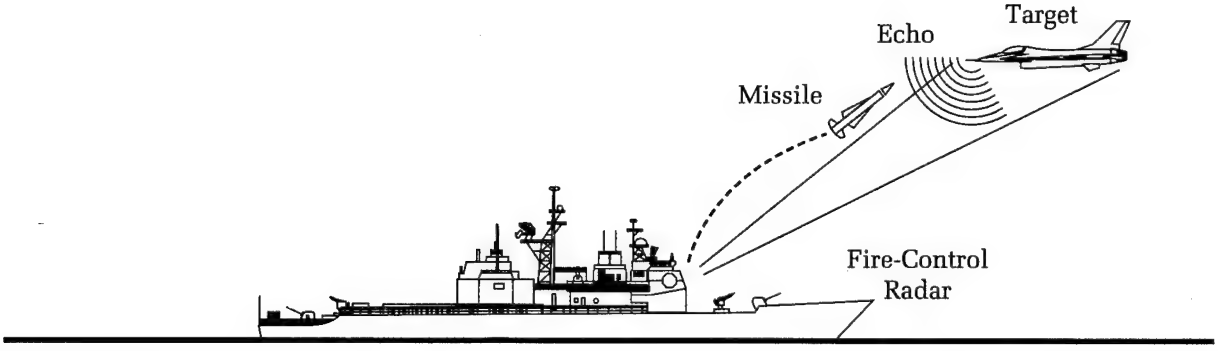
Option 1: Command-All-the-Way Guidance System – This weapon system includes a ship sensor for the target search, detect, and track function. The



Command-all-the-way guidance option.



Semiactive home-all-the-way guidance option.



Combined command mid-course/semiactive terminal homing guidance option.

Figure 1. *Missile guidance options.*

ship track sensor is responsible for acquiring the missile shortly after launch and tracking both the missile and the target during the engagement. The ship system then controls the missile through the entire flight by uplinking acceleration commands. For this option, the missile does not require an expensive onboard seeker.

Option 2: Semiactive Home-All-the-Way Guidance System – In this option, the target is illuminated in radio frequency (RF) by the ship. The missile acquires the target with a semiactive RF seeker shortly after its launch. The missile seeker then guides the missile throughout the flight.

Option 3: Combined Command Midcourse/Terminal Homing Guidance System – For this option, the missile flight consists of two phases: the midcourse phase and the terminal phase. During the midcourse phase, the missile is command guided, with the ship system performing the search, detect, and track functions. Similar to option 1, the ship track sensor acquires the missile after launch and tracks both it and the target during the midcourse phase. The ship system then controls the missile through the midcourse phase of flight by uplinking acceleration commands. Shortly before intercept, the missile seeker acquires the target, and the missile commences terminal homing. The missile seeker may be passive infrared (IR), semiactive IR laser, semiactive RF, or active RF. The ship sensor has the same functions as in the command-all-the-way option, but using a terminal homing phase lessens the accuracy requirements at the longer range intercepts.[1] The use of a semiactive missile seeker places an additional requirement on the ship, in that it must illuminate the target and command the missile seeker to be aimed in the appropriate direction for terminal acquisition.

This article examines the scope of the short-range AAW problem through the analysis of a defensive missile weapon system by first examining problems common to all three design options. The article then concentrates on the technical issues associated with RF and IR terminal homing missiles.

Technical Challenges Common to All Short-Range AAW Missile Systems

Some technical aspects of missile system engineering are common to all three of the missile system design options discussed in the Introduction. Many of these issues are related to the missile kinematics, including average speed, the missile time constant, and airframe maneuverability. The limited ship-sensor detection ranges in short-range AAW require the weapon system to provide all-aspect coverage with very short reaction times. For engagements at low altitude, the missile must possess high maneuverability to counter postulated threats. The implications of these requirements will be discussed in this section.

The Need for All-Azimuth Coverage

Vertical launching systems can offer a tremendous advantage over trainable launch systems for missile-based, short-range AAW systems. New missile weapon system designs are driven toward the use of vertical launch, because it allows each missile all-azimuth coverage. This capability is especially important for tightly packed raids coming from several directions. In ships with trainable launchers, the superstructure eliminates some launch azimuths, resulting in no-fire or cut-out zones. To achieve all-aspect coverage, several trainable launchers are required to be positioned around the ship. A vertically launched missile can simply delay pitching over until the ship superstructure is cleared (at some risk of loss of close-range effectiveness). Should one of the trainable launchers fail (through a mechanical failure or battle damage) or exhaust its missile supply, an incoming threat raid in its zone of fire could not be engaged even if the other launchers contain missiles and are healthy.

While vertical launch is preferable for short-range AAW missile systems, it does have some disadvantages. One disadvantage is a flight-time penalty that is greatest against low-altitude targets. A missile launched from a trainable launcher is already at the desired launch angle. The vertically launched missile ends up flying to a higher altitude during its pitchover maneuver. Therefore, the missile from the trainable launch system has a greater ground speed toward the predicted intercept point (Figure 2). For a given target range at missile launch, the flight-time penalty results in an intercept at a later time and at a closer position to the firing ship.[2] The lower average ground speed requires that the ship fire-control sensor detect the target at a longer range to maintain the same keep-out range. If the sensor is horizon-limited for a low-altitude threat, then the corresponding keep-out zone will decrease.

Figure 3 shows the advantage a trainable launch missile system has over a vertical launch missile system for minimum-range, low-altitude intercepts against a radially inbound target. Performance data for a

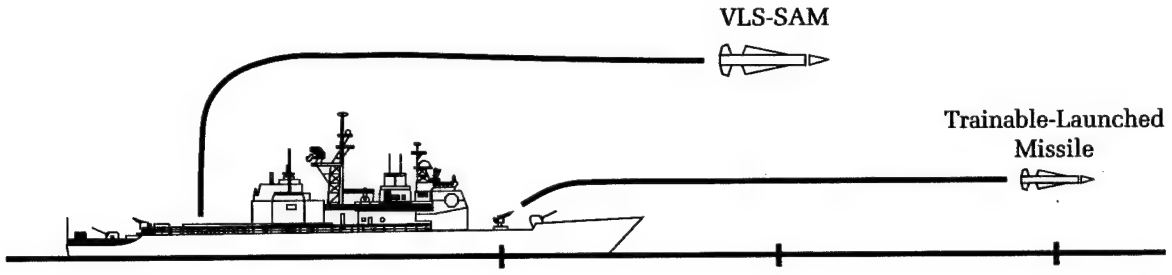


Figure 2. *Flyout ground speed toward target for vertically launched SAM and munition launched from trainable launcher.*

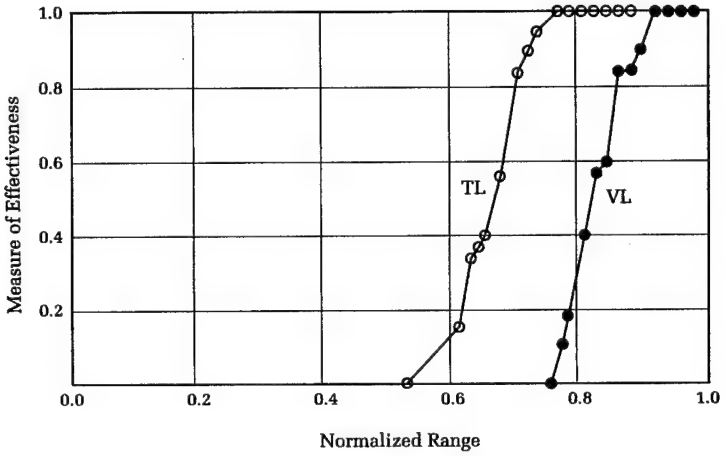


Figure 3. *The relative performance of a short-range AAW missile design using a trainable launch system and a vertical launch system.*

proposed short-range missile airframe, using a normalized measure of effectiveness, are presented as a function of the normalized intercept range. The data for the trainable launch system is designated *TL*, while the data associated with the vertical launch system is denoted *VL*. The results indicate a 20 percent reduction in minimum allowable intercept range. This can be an important consideration in detection-limited systems.

The physical control mechanisms for a vertically launched missile fall into two types: fin aerodynamic control and thrust vector control (TVC). The use of aerodynamic fin control provides a simpler mechanical design with greater reliability, while the TVC provides much greater missile responsiveness.

The TVC system is able to pitch the missile over considerably faster than aerodynamic fin control. This is possible because of a TVC system's ability to control the airframe to a much larger angle of attack. This control authority allows the TVC system to compensate for much of the time-of-flight increase associated with the vertical launch,[2] as illustrated in Figure 4. Short-range AAW missile systems are driven to use vertical launch to provide robust all-aspect coverage; TVC is required to optimize minimum-range performance. Therefore, in order to maintain battlespace in a short-range AAW environment, missiles using vertical launch systems are driven to incorporate TVC.

Limited Reaction Time

The need for maintaining the target intercept at a specified range from the ship is important for a number of reasons. These reasons may reflect the minimum effective range of the missile, the desire to conduct a target kill assessment in order to use a shoot-look-shoot fire-control policy, or the requirement of preserving a keep-out zone around the ship. The range at which intercept occurs is a function of: (1) detection range, (2) the interceptor's average speed to intercept, (3) the weapon-system reaction time from initial target detection to missile launch, and (4) target speed. The interplay of these elements on the intercept range can be approximated by making a number of simplifying assumptions. For this analysis, the missile is assumed to be flying at a constant average speed, V_m, at the same altitude as the target. The target is assumed to be flying

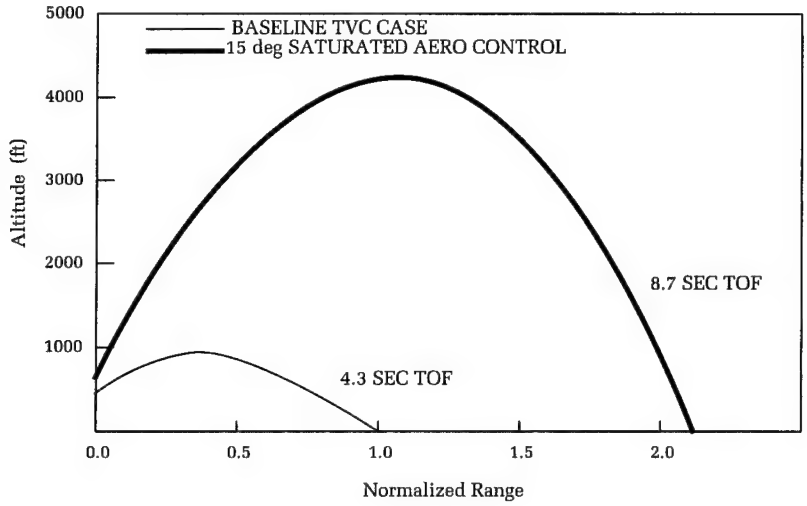


Figure 4. *Effects of thrust vector control on vertical launch pitchover.*

an inbound course at speed V_T. Detection of the target first occurs at a range R_D, with the ship weapon system having a reaction time of T_R before missile launch occurs. The intercept range, R_{INT}, can be expressed by:

$$R_{INT} = \frac{V_m(R_D - V_T T_R)}{V_m + V_T} \qquad (1)$$

Detection ranges of ship-based sensors are limited by the radar horizon, especially for low-altitude, sea-skimming targets, and the radar cross section (RCS) of the target. Therefore, a ship self-defense weapon system is required to have a short reaction time and a high-average-speed missile in order to maximize intercept range.

To achieve a high average speed over the short flyout ranges typical in ship self-defense scenarios, an interceptor missile must possess a high-boost acceleration capability. Thus, the missile design would require a high-thrust boost motor. However, the mass flow rate of the rocket motor design may conflict with the maximum flow rate constraints on a vertical launch system. A pulse rocket motor design, which uses energy management, can provide a balance between the need for high-boost accelerations at short-range intercepts with the need to maintain higher average missile speeds at longer intercept ranges. In a short-range intercept case, a high missile speed can be obtained by firing a second pulse shortly after the first pulse. For a longer range mission, a high average missile speed can be maintained by firing the second pulse after an optimized time delay. The timing of the second pulse minimizes the speed losses attributed to drag.[3]

In addition to maximizing the flyout speed of the missile, efforts can also be applied toward minimizing the reaction time of the weapon system. While a major portion of the reaction time is driven by the sensor system and the command and decision system (the time frame from track to missile-system receipt of an engage order), the missile-system design also contributes to the reaction time. Design factors that contribute to the overall reaction time include: (1) the settling time of the target track filters, (2) the time to declare firm target track, (3) the time for transferring alignment to the missile navigation systems, (4) the time for initiating missile components and firing thermal batteries, (5) the time for transferring power from the launch system to the missile round, and (6) the launcher sequencing time. Reductions in the system reaction time can be made possible by improving the target track filter transient response, launching missiles before firm track of the target is established, and developing missile components that require shorter warmup and initialization time. Solid-state electronics and quick-reaction thermal batteries are examples of advances in missile component technology.

Highly Maneuverable Targets

Many threats in the short-range AAW environment may be engaged at low altitude. The high atmospheric density at low altitude gives threats tremendous maneuvering capability. In this environment, missile miss-distance performance is driven by threat maneuvers instead of such factors as receiver noise, target glint, or heading errors,

which dominate at the higher intercept altitudes.

To illustrate the effects of a target maneuver, a simple constant amplitude sinusoidal weave will be considered. In such a maneuver, the target acceleration is given by:

$$a_T(t) = a_{T_{max}} \sin(\omega t) \qquad (2)$$

where

a_T is the target acceleration,
a_{Tmax} is the peak target acceleration (ft/sec^2), and
ω is the frequency of the target weave (rad/sec).

To continue this example, a simplified model of the interceptor dynamics flying against this sinusoidal target maneuver is examined. For this example, the interceptor is modeled as a single-lag system with the following assumptions: (1) the missile has an unlimited acceleration capability, (2) there are at least 10 missile time constants of homing time, and (3) the missile is assumed to be on an initial collision course with the target. A closed-form solution for the root-mean-square (RMS) miss is available for a missile using proportional navigation guidance, with an effective navigation ratio of N′ equal to 3. (The effective navigation ratio, N′, is a parameter that characterizes the guidance law response to various errors. Typical values for N′ are between 3 and 6.) The solution[4] is given by:

$$RMS\ \ MISS = \frac{a_{T\,max}}{\omega^2}\left(\frac{(\omega\tau)^3}{\sqrt{2}[1+(\omega\tau)^2]^{\frac{3}{2}}}\right) \qquad (3)$$

where

τ is the total missile time constant (sec), which includes the autopilot/airframe, noise filter, and seeker-loop time constants.

While the closed-form solution involves a number of simplifications, it illustrates the sensitivity of the RMS miss to the maximum target acceleration during the weave maneuver. For the linear model assumed, the miss is proportional to the maximum target acceleration. This equation illustrates the complex interrelationship of the interceptor total time constant with the frequency of the target weave. From the equation, there exists a target-weave frequency that maximizes the RMS miss. Typically, this maximizing frequency is of moderate value. At very high-weave frequencies, the miss rolls off because the maximum displacement of the target about the center path decreases to small values.

This analysis assumes that the interceptor is not acceleration limited. However, for targets capable of high accelerations, the effect of limiting will be important. Closed-form solutions do not exist for the acceleration-limited interceptor cases, so a numerical example is presented. As with the previous example, the interceptor is modeled as a single-lag system. The following assumptions are made: (1) at least 10 missile time constants of homing time are assumed, (2) the missile and target are assumed to be on a collision course, and (3) the missile uses proportional navigation guidance, with an effective navigation ratio of N′ equal to 4. Figure 5 presents the RMS miss, normalized by the maximum target displacement, as a function of the interceptor time constant, which is normalized by the weave period.[5] A separate curve is generated for each value of the interceptor-to-target maximum lateral acceleration constant. As this ratio increases, the RMS miss due to weave decreases. This figure illustrates the sensitivity of the miss distance to the total missile time constant. Missile systems are often driven to use small time constants in order to meet miss distance requirements against postulated target maneuvers.

These examples illustrate several of the parameters that influence the miss against a weaving target. These parameters include the missile time constant relative to the target weave period, and the target's maximum acceleration capability relative to the interceptor's maximum acceleration capability. Other effects that influence miss distance performance include variations in the proportional navigation ratio, N′, and a missile having an insufficient number of time constants of homing.[5]

Minimizing the interceptor's total time constant reduces the miss due to a target weave. This is generally true for other target maneuvers. Therefore, it is important to minimize the overall missile time constant. Missile autopilot designs with fast time constants have bandwidths that may encroach into the airframe bending modes. Using advanced structures technology to raise the airframe bending modes allows the development of faster airframe/autopilot time constants. Examples of such technologies include the use of composite materials to increase the missile airframe stiffness and the application of synthetic damping with magnetostrictive actuators.[6] Advanced control system methodologies offer the potential for more robust, less conservative autopilot designs that meet the required performance

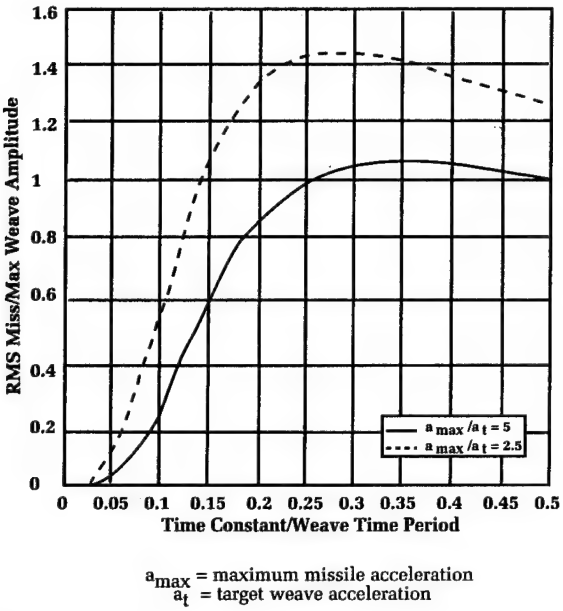


Figure 5. *Miss against maneuvering targets.*

specifications, while maintaining less separation between the first body mode and the autopilot bandwidth.[7,8]

Engaging highly maneuverable targets can result in large angles between the interceptor and target body axes and their velocity vectors, possibly degrading the fuzing process and warhead effectiveness. The main function of the proximity fuze, or target detection device, during the final stages of a missile's terminal homing phase is to adaptively initiate detonation of the warhead within the *lethal burst interval.* The lethal burst interval is defined to be that interval along a missile's trajectory in which the warhead can be detonated and inflict a maximum level of damage on the target's vulnerable areas. In order to balance the fuze and warhead performance of a missile for different targets and environments, the weapon system design process requires methods for determining the best design combination. Therefore, multidisciplinary models are used by weapon system engineers and designers to achieve a proper balance that is both practical and affordable. The incorporation of advanced computer technologies in future missile designs offers the potential for optimizing missile end game performance using onboard optimization techniques.

16 Technical Challenges for Short-Range AAW Missile Systems with Semiactive RF Terminal Homing

This section briefly discusses the technical issues associated with semiactive RF terminal homing missiles that use a monopulse seeker angle-tracking technique. This methodology could be applied to both the home-all-the-way and command/terminal homing missile guidance system design options. The use of monopulse receiver technology has greatly reduced the susceptibility to electronic countermeasures (ECM), but there is one major factor that may degrade angle tracking and terminal homing accuracy—signal multipath. RF signal multipath results from sea-surface reflections or the presence of multiple targets in a single raid. For a missile using RF terminal homing guidance, the seeker must discriminate the target signal from background noise in its field of view. The range at which target acquisition occurs determines the missile's homing time. Insufficient homing times can lead to unacceptable miss distance performance.

To facilitate an understanding of the effects of seeker angle tracking on missile guidance, a brief description of a typical system is presented. After acquiring the target, the missile begins its terminal homing phase using monopulse seeker angle tracking. The receiver's signal processor computes an estimated off-boresight angle, which the seeker track loop tries to null. The filtered seeker head rates are then used by the missile's guidance computer.

A typical monopulse seeker angle-tracking algorithm with doppler processing is now described in detail. A Fast-Fourier Transform (FFT) algorithm is used to estimate the power spectrum of the received total signal in terms of the doppler frequency shift relative to the rear reference signal. The FFT produces an estimate of the power in each doppler frequency cell based on a temporal sample set. Narrowband doppler filtering is applied to both the sum and difference channel signals. The individual targets or groups of targets in each cell may be considered to be individual tracks, and track files may then be maintained. Of all the doppler frequency cells, one must be selected as the *designated cell,* which is considered to be the best estimate of the cell containing the actual target. In one approach, the cell with the greatest sum channel power is selected. In another approach, any cell with a power greater than a specified threshold is qualified as a possible target. Of this set, the cell nearest the predicted target doppler frequency (uplinked from the ship system) or the cell that contains a target whose measured angular position is nearest the predicted value may be selected as the *designated cell.*

Only the RF signals in the designated cell pass through the doppler discrimination.

These signals are then used in the monopulse seeker angle tracking signal processing. The measured off-boresight angle, ε_m, at which the seeker measures the perceived target, may be computed by taking the real component of the ratio of the weighted difference channel signal to the sum channel signal for one type of monopulse system. This calculation may be approximated for a set of n equal strength sources by the equation:

$$\varepsilon_m = Re\left(\frac{\sum_{i=1}^{n}\varepsilon_i a_i e^{j\phi_i}}{\sum_{i=1}^{n} a_i e_i^{j\phi_i}}\right) \qquad (4)$$

where

ε_i is the boresight angle to the individual signal source,
a_i is the sum channel gain aimed at the individual signal source, and
ϕ_i is the relative phase angle of each signal source.

Because of interference effects, the sum channel signal can become small and, as a result, precautions must be taken in the design of the signal-processing algorithm in order to prevent the seeker track loop from being erratically driven by signal fades or noise.

Multipath Effects on Semiactive RF Missile Systems

When semiactive RF missiles engage targets at low altitudes, multipath effects present a significant problem. Signal multipath refers to the situation in which two or more electromagnetic waves emanate from a single source and propagate along separate transmission paths to a single receiving point. At low altitudes, signal multipath is caused by the sea-surface reflection of RF energy. A perfectly smooth sea surface would reflect RF energy at a single reflection point, with a single phase shift and a single attenuation value, both of which are functions of the signal's grazing angle. (The grazing angle is defined as the angle of incidence at which a transmitted signal strikes the sea surface and is reflected.) For an agitated sea-surface condition, a transmitted signal is diffused over an extended area of the sea. Many small, randomly phased reflections combine to form an incoherent composite reflection or signal. Therefore, a single receiving point *senses* several reflected signals that vary in phase and doppler frequency.

In a semiactive RF missile system, there are two dominant types of signal multipath: (1) forward and (2) back. Forward multipath refers to the existence of two signal transmission paths to a receiving point (i.e., the target). One transmission path involves the portion of illuminator energy that is reflected off the surface of the sea to the target. The second transmission path is the direct path from illuminator to target. Both of these signals are then reflected off the target and detected by the missile receiver.

A less important form of forward multipath is the arrival of multiple signals at the missile's rear reference antenna. Again, two transmission paths are involved. One path is the direct route from illuminator to missile. The other transmission path is the part of the rear reference signal that is reflected off the sea surface to the missile. In some instances, this form of multipath can degrade rear reference receiver lock and introduce frequency modulation noise into the missile's doppler processing.

Clutter is yet another form of forward multipath involving transmission from the illuminator to the missile main (or forward) antenna. This effect is important only for diffuse reflections over a relatively wide area. For the nearly antiparallel intercepts typical of short-range AAW, the signal arriving along this path may be removed by narrowband doppler filtering.

Back multipath, on the other hand, refers to the situation where the reflected illuminator energy from the target follows two paths back to the missile receiver. One path goes directly from the target to the missile receiver. The second path goes from the target to the sea surface, where it is reflected back to the interceptor.

The effects of the forward and back multipath on a missile receiver can now be examined for the case of pure specular reflection. For a perfectly flat sea surface, the combined signal at the missile receiver consists of RF energy at four distinct doppler frequencies. These frequencies are associated with each of the four possible transmission path combinations. The power spectrum of the signals is dependent on the reflection efficiency at the two sea-surface reflection points and the two reflections at the threat. The strength of the sea-surface reflection is a function of sea state, the signal's grazing angle, the illumination wavelength, and the illuminator polarization. Signal reflections are strongest at the lower sea states. One of the main consequences of the forward multipath return is that it can lead to fading of the illuminator signal at the target for near specular conditions.

Multipath returns degrade the angle tracking accuracy of monopulse receivers that

process a single sample of the sum and difference channels. It is possible that angular excursions of the seeker track loop can exceed the angular separation of the image and the target. The additional noise on the measured seeker line of sight (LOS) affects the missile guidance computer, leading to greater miss distances against low-altitude targets. The effect of signal multipath on the seeker's measured LOS is illustrated in the plot of the estimated and true off-boresight angles for a low-altitude engagement, as shown in Figure 6.

Misses attributed to signal multipath are distributed both above and below the true target altitude. At altitudes below 20 feet, the separation of the image and the target is small enough to have neglible impact on missile miss distance performance. For targets at altitudes above 200 feet, the reflected image is weak when compared to the real target return. Multipath effects are strongest for target altitudes between 20 and 200 feet.[9]

Methods for minimizing the effects of multipath returns at the missile receiver have been devised. The use of vertical polarization of the illumination signal can greatly reduce the effect of the sea surface return. Narrowband doppler filtering is used to selectively reject some of the return signals given sufficient frequency separation. However, in the case of the back multipath return, doppler filtering is able to distinguish between the reflected image and the target only a short time before intercept for specular reflections. For diffuse reflections, complete rejection of the multipath signals may not be possible.

Other signal processing techniques are available to aid in minimizing multipath effects. Super-resolution algorithms use different combinations of the monopulse quadrature antenna signals to achieve angular resolutions significantly smaller than the normal monopulse angular resolutions. This feature may allow angular resolution early enough in the homing phase to negate multipath effects. The use of dual-mode seekers can also minimize multipath effects by allowing signal information to be processed

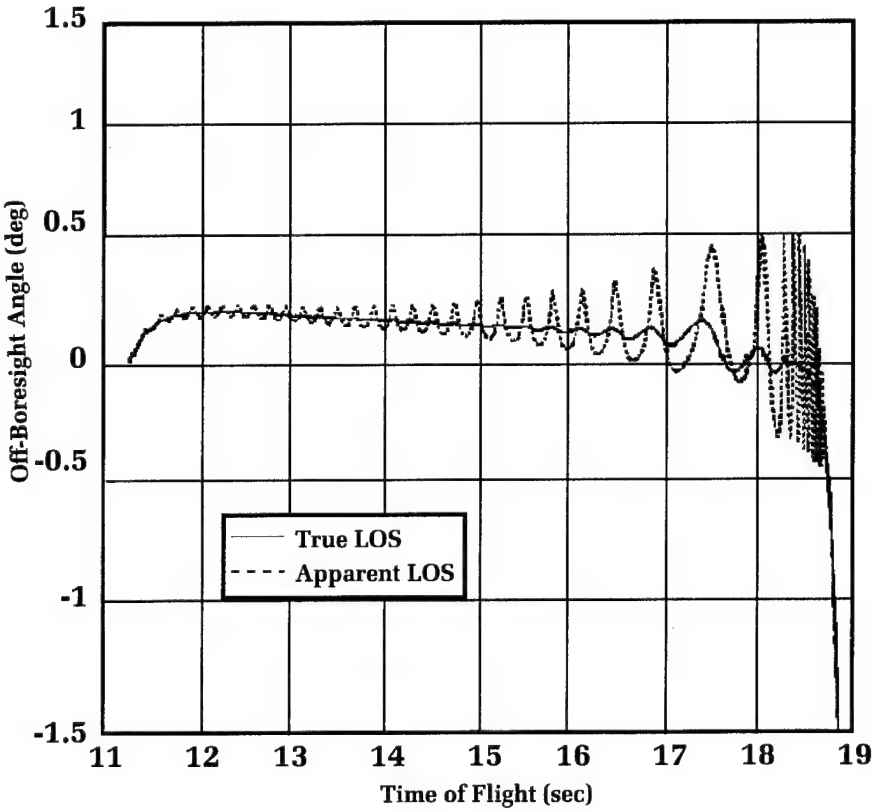


Figure 6. *Multipath effects on the measured line of sight.*

from two RF frequency bands. Because multipath effects vary with frequency, the returns may be combined synergistically to eliminate the multipath effects.

Another signal-processing technique, angle editing, refers to any method that alters the normal output of the error-angle processor in an attempt to reduce angle errors. The simplest form of angle editing is to exclude any deviant angle data from the track file. In this method, however, establishment of the correct boresight angle can be unreliable. This is because the phase distortions of the angular measurement are coincident with the back multipath signal fades. Angle editing techniques may be improved by eliminating only the angular excursions during the signal fades. This would allow the signal processor to more heavily weight the angular measurements outside of the signal fades.

Trajectory shaping methods may also be used to overcome multipath effects. In one method, a missile's trajectory can be shaped in the vertical plane to permit the missile to approach the target and the sea surface at a smaller flight-path angle during the terminal homing phase. This shaping may be achieved by modifying the final flight-path angle in the missile's guidance law. The trajectory is designed so that the missile flies at a grazing angle that minimizes the sea surface reflection coefficient at a preselected point in the homing period. The angle where the surface-reflection coefficient is at a minimum is referred to as Brewster's angle. In another trajectory-shaping procedure, the missile can be commanded to fly at a fixed altitude until a short time before intercept. This technique prevents multipath effects from corrupting elevation control of the missile during the period before resolution.

Multitarget Effects on Semiactive RF Systems

Several postulated threat scenarios against surface ships consist of low-altitude, high-speed targets flying in multiple vehicle raids. For a semiactive missile system using RF terminal homing, the presence of multiple targets can severely degrade monopulse seeker track accuracy. This degradation is caused by the relative phasing effects of the multiple signal returns.

This section looks at the unique aspects of the multitarget raid. The basic scenario at handover is shown in Figure 7. The definitions of several errors that were considered at handover are included in the figure. It is assumed that the ship system has resolved the raid and has designated the front target as the primary target for interception. The performance of the system in this section will be based on the measurement of the interceptor's miss distance relative to the primary target.

The following assumptions are made for this analysis: (1) handover to missile semiactive RF homing has been completed; (2) the missile has acquired the target; (3) the raid consists of two targets; (4) the targets fly in a fixed formation at an altitude of 30 feet with a constant speed of 3000 ft/s; (5) all heading errors, seeker-pointing errors, and miss distance values are measured relative to the primary target; (6) the missile has sufficient maneuverability to steer out heading errors of the same magnitude as expected against a single target raid; (7) RF multipath and clutter effects are neglected; and (8) the interceptor has a small diameter and large seeker beamwidth.

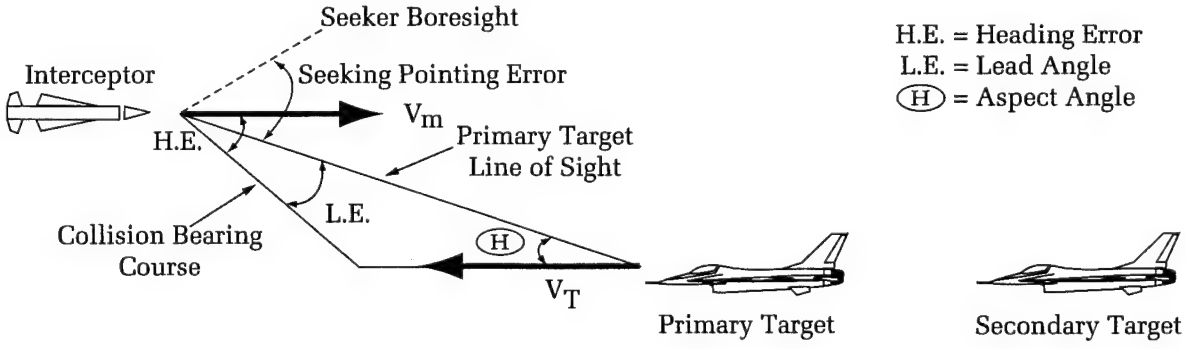


Figure 7. *Handover geometry for multitarget analysis.*

The engagement scenario is illustrated in Figure 8. As shown in the figure, the interceptor is diving at a flight-path angle of -20 degrees on a multiple-target stream attack. One parameter of interest is the horizontal spacing of the raid. Because of the geometry involved, the axis of target scintillation is readily visible to the interceptor. Therefore, large miss distances are to be expected, even for perfect handover conditions.

The perfect handover results are presented in Figure 9. Miss distances are significant for all target separations and become worse as the separation between the two targets increases. Since the axis of scintillation is not normal to the interceptor's LOS to the target raid, the larger target separations do not result in zero misses. This is because the angular extent of the raid does not exceed the interceptor beamwidth.

The effects of heading and seeker pointing errors are shown in Figure 10. Neither has a strong effect on the miss distance. Therefore, improving the handover accuracy does not alleviate the unique miss distance effects of the multitarget problem.

Proper design of the interceptor's midcourse trajectory shape can create the best aspect, or approach angle, of the interceptor relative to the target raid, as defined in Figure 7. The effect of the chosen approach angle on the two-target stream attack is considered. The flight-path angle at handover is determined by the heading error and the relative missile-target geometry due to the specified aspect angle at handover.

Figure 11 shows the results against the 100-foot spaced stream attack. In this plot, the miss is shown as a function of the specified handover aspect angle and the heading error. Clearly, from these results, the miss is a strong monotonic function of the handover aspect angle. This indicates that it is best for the interceptor to attack in the direction parallel to the scintillation axis of the raid. For crossing targets, the implication of this strategy is that azimuthal trajectory shaping can be used to bring the interceptor antiparallel to the raid axis at handover and significantly improve terminal performance.

The effects of selecting the noise-filter time constant, the seeker-head track-loop time constant, and the total missile time constant are now considered. Studies have indicated that the miss due to multiple targets is related to the total missile time constant. For a stream attack with 100-foot target separation, the interceptor miss distance was minimized using a total time constant of approximately 1.0 second. Other raid spacings may change the value of the minimizing time constant. The selection of the missile time constant must also include consideration of target maneuver, receiver noise, and target glint effects.

Reduced Target RCS and Reduced Homing Time

Future antiship cruise missiles may have reduced RCS designs. In a short-range AAW scenario, a reduced target RCS affects the acquisition range of the missile seeker in both semiactive RF and active RF weapon systems. The reduced acquisition ranges can lead to intercepts with the missile having insufficient homing times. This situation would produce large miss distances. It has been shown that at least 7 to 10 missile total time constants of homing are required to

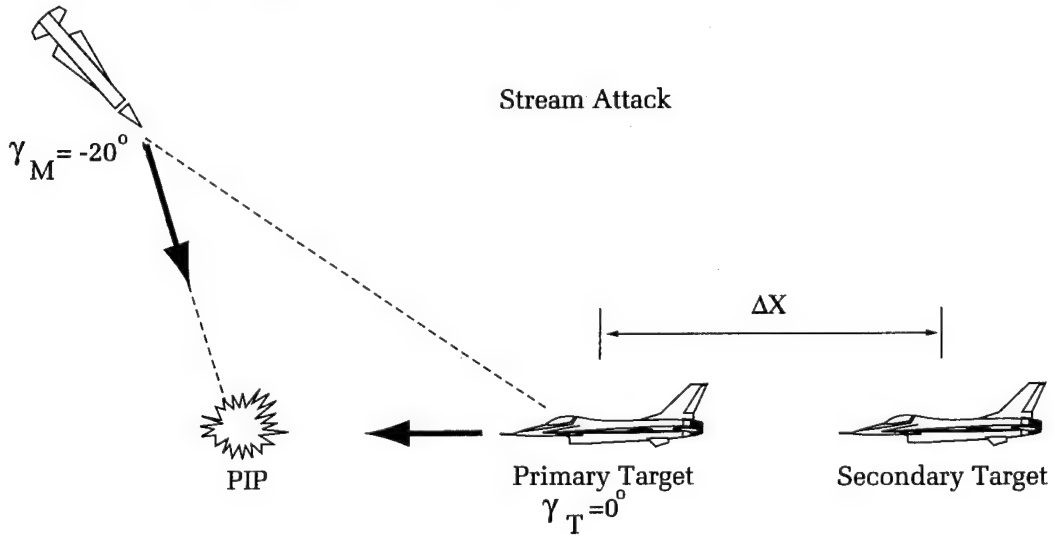


Figure 8. *Example short-range AAW multitarget scenario.*

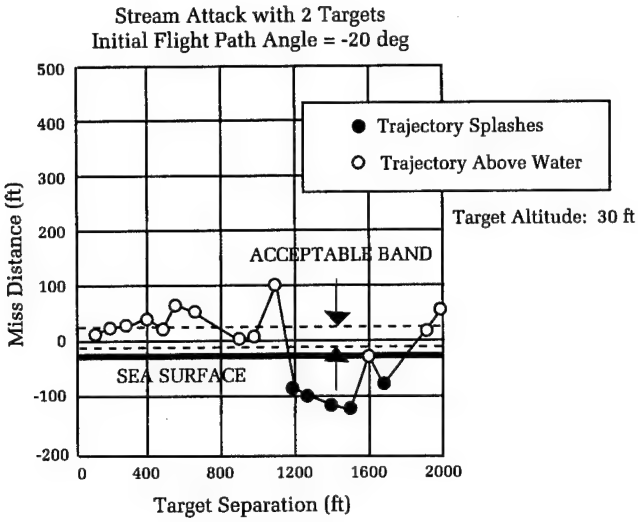


Figure 9. *Miss relative to the primary target for perfect handovers.*

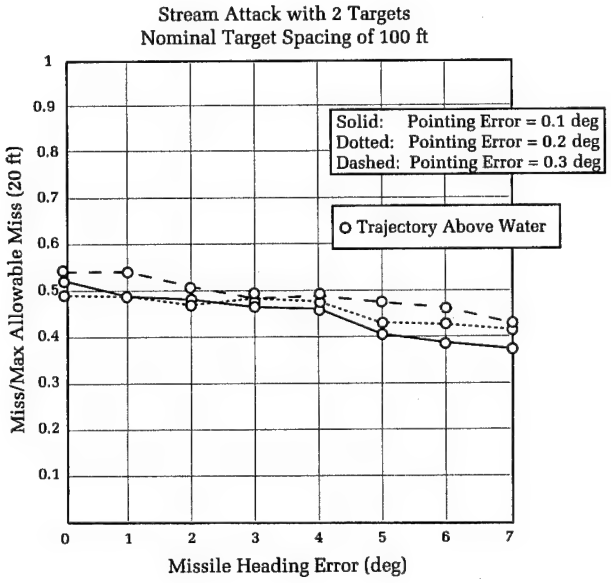


Figure 10. *Effect of handover errors on miss against multiple targets.*

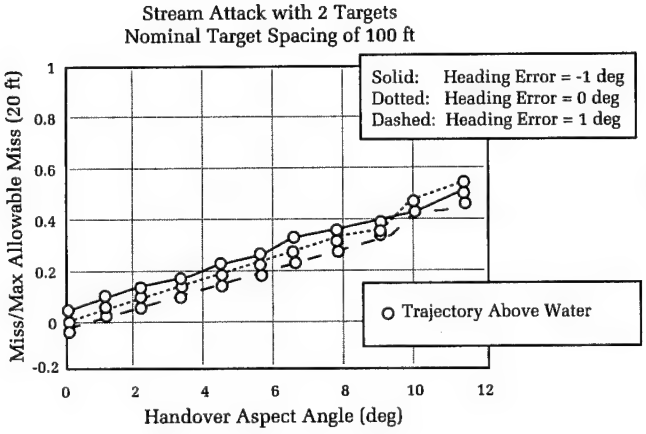


Figure 11. *Terminal aspect angle effects on miss against multiple targets.*

steer out heading errors and target maneuvers.[10]

The effect of reducing the RCS of a target is shown in Figure 12. This figure presents two engagement envelopes (shown as a function of normalized range and altitude) of a conceptual missile airframe. Ship system limitations and multipath effects were neglected in the simulation analysis. Both envelopes were constructed from missile performance data generated by a six degree-of-freedom (DOF) simulation. One engagement envelope illustrates the ability of the missile design to obtain a specified measure of effectiveness against a target with a nominal RCS. The other envelope shows the capability of the missile airframe design to intercept a target with a 20-dB reduction in its RCS for the same measure of effectiveness. The reduced RCS clearly has a significant effect on reducing the vertical plane engagement envelope.

Technical Challenges for Short-Range AAW Missile Systems with IR Terminal Homing

Some problems associated with a short-range missile using IR terminal homing are now discussed. Both the insufficient IR signatures of small, low-speed targets and the high atmospheric attenuation of IR signals during poor weather can lead to reduced acquisition ranges and shortened homing times for a defensive missile system. The presence of the sun in the IR seeker's field of view can also disrupt target acquisition and the angle tracking process. When IR homing missiles are launched in dual salvos, problems of fratricide may also develop. Each of these subjects are discussed in more detail later in this section.

A modern IR seeker uses focal plane array (FPA) technology consisting of an array of IR detectors. Each detector has a very small field of view, thereby giving the FPA IR seeker much finer angular resolution than an RF seeker. Whereas the RF seeker uses frequency discrimination, the IR seeker uses spatial discrimination. The basic processing function of the IR array is to compare each detector output to the outputs of surrounding detectors. In some systems, individual detectors are determined to be "on or off" by using a threshold value that may be adjusted in terms of the estimated environment and detector thermal noise to provide a constant false alarm rate. Various algorithms may use the thresholded pixel patterns to determine the *designated cluster.* In a gray scale system, the brightness of the various detectors may be used to identify the *designated element* or *designated cluster* that has the highest output ratio. Typically, only a subset of the FPA is under active consideration. A track or acquisition box may be defined about the expected position of the target. Any images outside of this box are not considered in tracking or acquisition. The estimates of background noise are typically analyzed only over the track box. In some cases, a slightly larger box may be defined for the collection of background noise statistics.

The movement of the designated element across the focal plane is used to drive the seeker head track loop, which attempts to maintain the target within the field of view by correcting for missile-to-target LOS variations and coupled body motions. The LOS rate typically required for terminal guidance may be reconstructed from the movement of the designated element (which is equivalent to the boresight angle rates) and the seeker head rates.

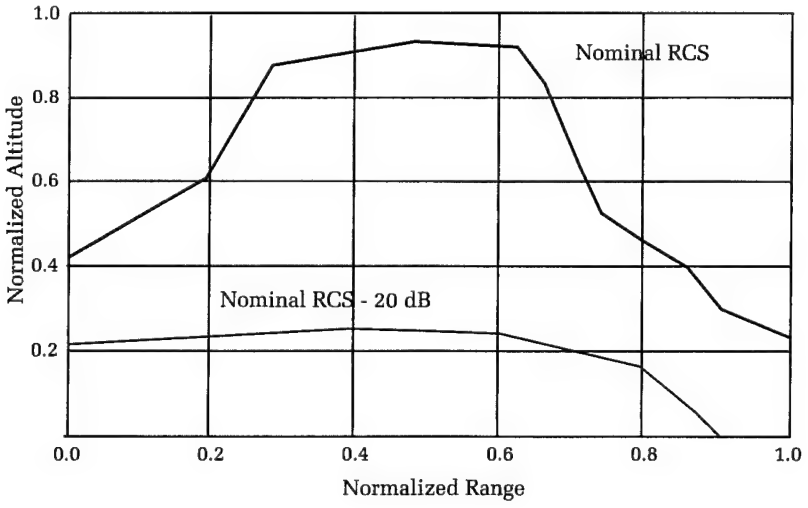


Figure 12. *Effect of reduced target radar cross section on a vertical plane engagement envelope.*

Low Visibility and Sun Corridors for IR Systems

For missiles with passive IR seekers, acquisition of low-speed targets with small IR signatures or targets with the sun in the seeker's field of view is difficult. Up to now, trajectory-shaping concepts for an interceptor missile have been concentrated only on the vertical plane. For radially inbound targets at low altitudes, vertical-plane trajectory shaping may cause the target to be seen in a near head-on geometry. When targets are viewed head-on, they present small areas to the missile's seeker and, thus, may have small IR signatures. Therefore, a radially inbound target may be difficult to acquire, resulting in the missile having insufficient homing time. Figure 13 shows the horizontal plane engagement envelope of a missile design using IR terminal homing against a low-speed, low-altitude target.[11,12] The darkened regions in the figure represent successful intercepts; reasons for missed intercepts are indicated by a symbol. The *HT* indicates a miss due to insufficient homing time resulting from inadequate acquisition range. The difficulty in acquiring the radially inbound target is clearly demonstrated. For crossing targets, seeker acquisition occurs sooner because of the improved aspect angle of the missile relative to the target.

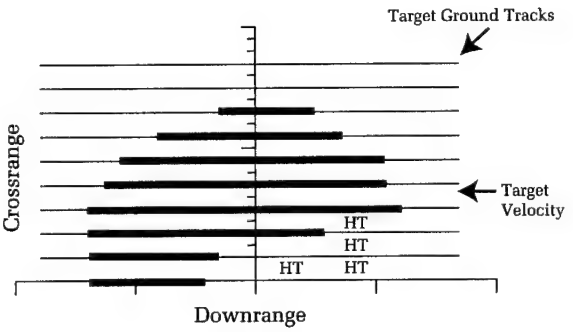


Figure 13. *Horizontal engagement for a short-range AAW missile using IR terminal guidance against a low-speed target.*

Studies performed for the Surface-Launched Weapons Block have shown that the use of three-dimensional trajectory shaping can improve the acquisition performance of missiles with passive IR seekers, and improve the overall system probability of hit. Interceptor trajectories against a radially inbound target, as shown in Figure 14, can be optimized by using an NSWCDD-developed Multidisciplinary Trajectory Optimization Program to give a spectrum of specified aspect angles relative to the target. The trajectory and aspect angle that maximizes overall missile system performance can be determined by using this approach. Passive IR homing system performance against low-speed targets can be improved by the use of combined vertical/azimuthal trajectory shaping. Azimuthal trajectory shaping can also be used to place the sun outside the IR seeker's field of view.

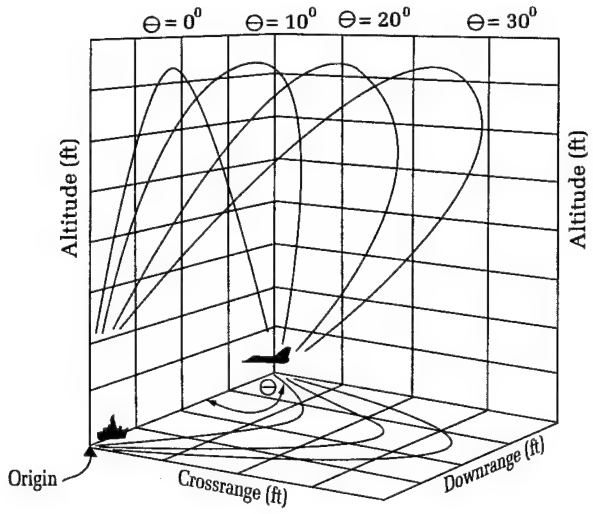
All-Weather Capability Considerations

The presence of atmospheric water vapor can be significant at low altitude. During periods of bad weather, clouds, fog, rain, and snow, severe attenuation of a target's IR signature may occur. This can greatly reduce the acquisition range of the IR seeker and lead to very short acquisition ranges and inadequate missile homing times. Because of this effect, new missiles are being designed to use both RF and IR homing. Examples include the Standard Missile SM-2 Block IIIB and the Sea Sparrow RIM-7P missiles. The performance of RF seekers is not degraded as severely in bad weather when compared to IR systems. Therefore, a multisensor missile system would be able to use the RF seeker for terminal guidance when the IR seeker is not able to detect the target because of weather-related performance degradation. The benefits of the IR seeker could be exploited during fair weather, while the RF seeker would provide a graceful degradation in overall missile performance during poor weather. It should be noted that using IR homing will be advantageous in a high percentage of the maritime environments to be encountered in the future littoral warfare scenarios.

Interceptor Fratricide

Limited engagement times restrict the available depth-of-fire for a short-range AAW missile weapon system. (The depth-of-fire is defined to be the number of sequential rounds or engagements that are necessary to achieve a given level of kill probability for a single target.) In order to achieve the required probability of kill against a raid and to overcome the less than perfect effectiveness of ordnance, a dual-salvo firing policy is required. Fratricide between the dual-salvo missiles presents a dilemma, particularly for missiles using IR terminal homing.

Two possibilities of fratricide can occur for IR homing missiles. The first possibility



Θ = Trajectory Offset Angle

Figure 14. *Trajectories with azimuthal shaping to improve IR seeker performance.*

may occur when the second missile of the dual-salvo acquires and then homes on the first missile of the salvo instead of the target. Together, the aerodynamic and propulsive heating effects, and the proximity of the two interceptors, can create scenarios in which the first missile's IR signature may be stronger than the target(s). For a typical low-altitude intercept, both the target and the first missile of the salvo may fall within the second missile's seeker field of view shortly before handover. If the two images are in proximity during the IR seeker acquisition process, the acquisition algorithm may inadvertently declare the first missile to be the target, because it presents either the brightest image or the image that exceeds a threshold value nearest the weapon system's predicted target location. One approach for preventing inadvertent acquisition of the first missile is for the weapon control system to uplink a tactical map of the target and interceptor environment. The second missile would then conduct a seeker scan to generate a *snapshot* of the search region. After the scan is completed, target characteristics, such as signal strength, or LOS rates, would be used to determine the designated target. Development of dual-band FPAs may allow for the discrimination of the target's IR signature over the two separate bands (in effect *coloring* the image). The first missile's IR signal strength will not vary over time in the same manner as the actual target's IR signal because of the differences in closing velocities relative to the second missile. Thus, the history of the two signals over some time window could be used for discriminating between the two images and preventing declaration of the salvo's first missile as the target.

Once the second missile of the salvo has acquired the target, coupling of the IR seeker's fine angular resolution with track file-maintenance logic could prevent switching of the track declaration from the real target to the first interceptor. Such an approach requires making assumptions about the behavior of the target and the first missile. An example of a target characteristic assumption would be the expected motion between scene dwells. This information could be used to determine which one of the images on the FPA is the same image as was declared on the previous dwell.

For missile systems that do not possess IR signal processing capable of preventing the acquisition of the first missile in the salvo, NSWCDD analysis has indicated that azimuthal shaping of the second missile's trajectory effectively separates the first missile and the target so that the first missile is not within the second missile's seeker field of view at typical IR seeker start search times. The first missile can be launched along a nominal trajectory, while the second missile is launched at a shallow (less than 15 deg) azimuth off the bearing to the predicted intercept point. The second missile then steers toward the predicted intercept point and arrives with an azimuth bearing of the same magnitude but opposite polarity. The second missile's shallow azimuthal flyout angles provide sufficient separation of the salvo with small speed loss.

The other possibility of fratricide associated with dual salvos occurs when the smoke plume of the first missile obscures the second missile's view of the target. Depending

on motor exhaust characteristics and atmospheric conditions, the smoke plume may sufficiently attenuate the target's IR signature so that track loss occurs. Accurate prediction of plume expansion in the atmosphere is difficult, but preliminary NSWCDD analysis of a typical low-altitude intercept indicates that the LOS from the second missile to the target may pass through the first missile's exhaust plume for significant periods of time during the terminal homing phase. Precautions may be required in the rocket motor design or in the trajectory shaping to prevent the exhaust plume from degrading the homing performance of the second missile.

Summary and Conclusions

Some of the most important technical problems and issues associated with the short-range AAW environment have been presented. A matrix of missile-system design drivers, their implications, and design modifications with applicable technologies is presented in Table 1. In the authors' opinion, the most important concern involves the high-g target maneuver at supersonic speed and low altitude. Countering this threat requires further reductions in the missile time constant and improvements in missile structural technologies. This required reduction may conflict with the desired time constant that minimizes multitarget scintillation effects and even single-target RF glint-noise effects. In order to reduce the overall missile time constant, it may be necessary to apply advanced airframe control theory in the design of the missile autopilot. Required structural advances include developing airframes that can sustain higher maneuver loads, as well as stiffer airframes to maintain the first body-bending frequency outside the broader bandwidth of the autopilots.

Multitarget raid and signal multipath effects present serious problems for RF terminal homing missiles. Trajectory shaping can reduce multipath and, in some cases, reduce multitarget effects. The use of advanced signal-processing techniques, such as a super-resolution algorithm, may provide a solution to these problems

Limited ship sensor-detection range is an important restriction on short-range AAW performance and may be addressed by the application of vertical launch and TVC.

The application of FPA IR seeker technologies can increase angular resolution and solve many of the problems associated with multipath and multiple target raids; however, problems with low IR signature targets will remain. These problems can be alleviated by azimuthal trajectory shaping. IR technology also has the limitation of not having an all-weather capability. Hence, multimode RF/IR sensor integration in missile systems is being examined.

If sufficient improvements in ship sensor track accuracies can be obtained, then the capability of the ship sensor may exceed that possible with an onboard semiactive or passive missile seeker. This would allow for the use of a pure command-guidance system over the entire AAW environment, with an associated reduction in the missile's flyaway cost. Another possible benefit of improved ship sensor-track accuracy may be realized through the optimal blending of ship track sensor and onboard missile sensor information during RF terminal guidance to improve homing performance in the presence of multipath and multitarget raids.

Trajectory shaping provides an inexpensive method to overcome several of the short-range AAW mission challenges. This approach has not been fully exploited in missile designs to date. The traditional trajectory design approach has been to maximize the final speed of the interceptor through proper energy management. A true systems-engineering approach to a missile's mission and trajectory design may yield trajectories that maximize the weapon system's overall performance by sacrificing missile speed for an associated increase in sensor performance.

Considerable potential remains for advances in short-range AAW missile technology by the proper application of advanced control systems, structures, sensors, and trajectory optimization techniques.

Acknowledgments

As a broad review article, the authors have touched on many people's efforts, as indicated by the references. Mr. Ernest J. Ohlmeyer, Mr. Danny L. Brunson, Dr. Thomas J. Rice, Mr. Thomas H. McCants, Jr., and Dr. Jacques E. Goeller provided a great deal of insight and useful commentary on this article. Additionally, Mr. Brian L. Kiser, Mr. David B. Hanger, Ms. June C. Drake, and Mr. Craig A. Martell provided a significant portion of the analysis presented in this article.

Table 1. *Matrix of Missile System Design Drivers*

Design Drivers	Impact/ Design Requirement	Design Modification/ Applicable Technologies
High-speed ASM	Short reaction times/ Quick-response system	Solid-state electronics Fast batteries
	Limited Battlespace/ High-average-speed missile	High-thrust boost rocket motors Pulse motor technology
Highly Maneuverable ASM	Missile maneuverability/ Small missile time ratio	High g-frame (Advanced structures) Fast autopilot Modern control theory (H-infinity, μ-analysis) High performance control surface actuators Stiff/active airframe
Sea-skimming ASM	Limited ship system detection ranges	Multisensor (RF/IR) integration for improved ship system detection
	RF signal multipath effects at low altitude/Image rejection	Vertical-plane trajectory shaping Vertical polarization of illumination Advanced signal-processing techniques
Multiple ASM Raids	Target discrimination	Seeker super-resolution methods Dual-mode RF guidance
	Interceptor fratricide	Three-dimensional trajectory shaping Advanced signal-processing techniques
Reduced-Signature ASM	Reduced detection and acquisition ranges Reduced homing times/ Small missile time ratio	Multisensor (RF/IR) integration on missile Fast autopilot (see above)
Environmental Factors	Weather effects/All-weather capability Sun corridor effects for IR systems	Multisensor (RF/IR) missile guidance Three-dimensional trajectory shaping

Missile Guidance and Control Term Definitions

The concept of missile guidance and control can be defined by analyzing the terms, **guidance** and **control**. The first term, **guidance**, refers to the process of tracking, computing, and directing a missile to a designated target. A missile guidance system consists of components both internal and external to the flight vehicle. These components are responsible for directing the missile onto a desired flight path from the launch platform to the target. The second term, **control**, refers to the process of steering the missile during flight. Components of the control system, located within/on the missile airframe, are responsible for maintaining the missile in a proper flight attitude in relation to the desired flight path. Collectively, a missile's guidance and control systems determine the success of the mission against a target.

Missile guidance is generally divided along the three phases of a missile's flight:

1. Boost
2. Midcourse
3. Terminal

The boost phase begins with missile launch and usually ends at booster burnout, or when the missile is at the desired flight speed. In the case of a vertically launched missile, the boost phase is considered to be the time period it takes the missile to pitch over to a desired flight-path angle. The objective of the boost phase is to position the missile at a point where it can either acquire the target or receive external guidance commands. The midcourse phase of missile flight is often the longest in terms of distance and time. The objective of the midcourse phase is to place the missile on the desired course to the target. During long-range flights, trajectory shaping is performed in the midcourse phase to maintain speed and position the missile near the target for the terminal phase of flight. The terminal phase is the final segment of a missile's flight. It is in this phase that the last flight-path corrections are made to keep the missile on a collision course with the target. The performance of the guidance system is the most important element of the missile design in the terminal phase. The different options available to a missile designer are presented in the next paragraph.

Missile guidance systems are classified into four categories. They are:

1. Preset guidance
2. Command guidance
3. Beam rider guidance
4. Homing guidance

For the purposes of this article, only two of these categories, command guidance and homing guidance, are discussed.

Command Guidance

In command guidance, the flight path of the missile is directly controlled at all times by the launching platform. All steering commands are transmitted to the missile during flight via the fire-control radar beam, radio command link or, as in the case of the TOW antitank weapon system, through wire. In one form of command guidance, command to line-of-sight (CLOS), the launch platform keeps the missile directed along the imaginary line to the target, or LOS.

The overall performance of a command-guided missile is determined by the quality of the launch platform's target illumination and receiving system. One advantage of command guidance is that the missile does not require an expensive onboard seeker. However, a major disadvantage of a command guided missile is that it is vulnerable to jamming and ECM.

Homing Guidance

Homing guidance represents the most accurate method of guiding a missile to a target. A missile system that uses homing guidance relies on the target to be the source of guidance error signals. There are three types of homing guidance. They are:

1. Active homing
2. Semiactive homing
3. Passive homing

Each of the types refers to how the missile acquires information from the target.

Active Homing

A missile using active homing guidance has both a transmitter, which generates the target illumination signal, and a receiver, which senses the energy reflected from the target. The reflected target energy is processed by the onboard flight-control system to provide steering commands during the terminal phase of flight. The illumination signal can be in a variety of forms: RF, millimeter wave, or laser energy. Phoenix, Harpoon, and AMRAAM are examples of active homing missiles.

There are two advantages to active homing guidance. First, a weapon system that uses active homing missiles has a fire-and-forget capability. This would allow the launch platform to engage different targets in a short period of time. Second, active homing guidance allows for over-the-horizon targeting.

There are also disadvantages to active homing missiles. The most important one is that the weight and power limitations imposed on the transmitter design may restrict the missile seeker acquisition ranges. The need for both a transmitter and a receiver onboard the missile also increases the complexity and cost of the round. An active homing missile betrays its approach to a target. This would permit the target to make evasive maneuvers or employ ECM. Another disadvantage is that once fired, the launch platform has no further control over the missile. This prevents a missile from being directed to another target by the firing platform.

Semiactive Homing

In this form of homing guidance, the target illumination signal is supplied by another source, usually the launch platform. The reflected energy is received by a seeker onboard the missile and used to compute flight-path corrections in the missile's terminal phase. Semiactive homing guidance offers advantages to ship-based missile systems. One advantage is that it provides for the use of shipboard transmitters to illuminate the target, resulting in greater illumination power at short to medium intercept ranges. However, the long illuminator tie-up times for the medium range intercepts could be a possible drawback. Another advantage to semiactive homing is the fact that the missile does not have to carry a transmitter, resulting in more space within the missile for additional rocket motor propellant, ordnance, or guidance equipment. Examples of missiles employing semiactive homing guidance include the U.S. Army's Patriot missile and the U.S. Navy's Standard and Sea Sparrow missiles.

Passive Homing

A missile employing passive-homing guidance uses the latent energy emanating from the target to provide terminal-homing information. Target signals can be in the form of sound, heat, or RF radiation. RAM, Stinger, and Sidewinder are examples of missiles that use passive RF terminal homing guidance. The Standard ARM and the Shrike, two air-to-surface antiradiation homing missiles, use passive RF terminal homing. Passive homing guidance offers two advantages. One advantage is that, once launched, a passive homing missile does not require support from the launch platform. The second advantage is that a missile does not have to carry a transmitter, allowing space for additional ordnance, rocket motor propellant, or guidance and control equipment.

The principal surface-to-air missile system employed in the U.S. Navy today, the Standard missile, uses a combination of command and homing guidance techniques. During the midcourse phase of flight, the missile is command guided. Steering (acceleration) commands and seeker-pointing commands are computed by the weapon system's computer onboard the ship and then uplinked to the missile. Once the seeker acquires the designated target, the missile transitions into a semiactive RF terminal-homing phase.

References

1. Nesline, F. W. and Nesline, M. L., "Analysis of Optimal Command Guidance vs. Optimal Semiactive Homing Missile Guidance," *Proceedings of the 1986 American Control Conference*, Boston, MA, Jul 1985.
2. Solis, R., "An Analysis of the Vertical Launch Phase of a Missile Concept," *AIAA 21st Aerospace Sciences Meeting*, Reno, NV, 10-13 Jan 1983.
3. Phillips, C., "Energy Management for Multiple-Pulse Missiles," *Journal of Spacecraft and Rockets*, Vol. 27, No. 6, 1990, pp. 623-629.
4. Chadwick, W. R., *An Analysis of Collision-Course Navigation with Command Guidance*, Australian Defence Scientific Service Weapons Research Establishment, WRA Technical Note SAD 141, Dec 1964.
5. Drake, J., "Weaving Target Analysis," Aeromechanics Branch Technical Memorandum, Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA, Jan 1992.
6. Schindel, L. and Hardy, S., "Innovative Missile Airframe Concepts," *Naval Surface Warfare Center Technical Digest*, NSWC MP 91-44, Naval Surface Warfare Center Dahlgren Division, Silver Spring, MD, Sep 1991, pp. 84-91.
7. Ohlmeyer, E. J., *Robust Control Theory: Current Status and Future Trends*, NSWC MP 90-385, Naval Surface Warfare Center Dahlgren, Dahlgren, VA, Jun 1990.
8. Bibel, J. E. and Stalford, H., "An Improved Gain-Stabilized Mu-Controller For a Flexible Missile," *30th Aerospace Sciences Meeting and Exhibit*, Reno, NV, 6-9 Jan 1992.
9. Hagan, J. D., "Low Elevation Target Multipath Considerations for Semiactive RF Air Defense Missile Systems," Naval Surface Weapons Center, Dahlgren, VA, 1985.
10. Bibel, J. E., *Application of the Adjoint Method in Homing Missile Performance Analysis*, NSWC TR 88-311, Naval Surface Weapons Center, Dahlgren, VA, May 1990.
11. Christie, J. C. and Phillips, C. A., *Missile Homing Improvement Program (MHIP) Infrared (IR) Seeker Model*, NSWC MP 90-531, Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA, Sep 1990.
12. Phillips, C. and Kiser, B., *Surface-to-Air Missile Engagement Envelope Generator (EEG)*, NSWC TR 89-89, Naval Surface Warfare Center, Dahlgren, VA, Jul 1989.

The Authors

CRAIG A. PHILLIPS received a B.S. with Honors in Aerospace Engineering from the University of Texas at Austin and began his career in the Division in 1981. Selected for the Division's full-time advanced study program he earned an M.S. in Aerospace Engineering from Virginia Tech. He has worked extensively in the area of missile-system performance analysis, concept definition, and trajectory and propulsion optimization. Recently, Mr. Phillips served as principal investigator on an Independent Exploratory Development project concerning inertial navigation system (INS) and global positioning system (GPS) attitude sensor integration. He is a member of the American Institute of Aeronautics and Astronautics and serves as a reviewer for its *Journal of Guidance and Control.*

JAMES M. CHISHOLM graduated from the Pennsylvania State University in May 1983 with a B.S. degree in Aerospace Engineering. He has been employed at NSWCDD since September 1983. As an aerospace engineer with the Aeromechanics Branch of the Weapons Systems Department, Mr. Chisholm has worked extensively in the area of missile system simulation and performance analysis.

Integrated Interior Communications and Control—Engineering Validation of a Total Ship Architecture

Franklin A. Ervin

The Navy/Marine Corps "From the Sea" and the Joint Chiefs of Staff "C^4I for the Warrior" strategies recognize the need to balance reduced joint forces with increasingly flexible options to deploy them in combat. This requires a focus toward information fusion; tactical decision-making under stressful, quick-response situations; and development of robust communications methodologies to support adaptable total ship functions and missions.

This article describes an Integrated Interior Communications and Control $(IC)^2$ architecture that integrated the ship functional departments such as combat systems, operations, navigation, engineering, air, and administration on a command network to maximize connectivity, integrate data, and manage information. The development of the $(IC)^2$ architecture is traced through a proof-of-concept demonstration and an engineering validation conducted in 1991 and 1993, respectively. The outcome of these system engineering efforts was an evolution of the $(IC)^2$ architecture to a stage that is mature enough for design and construction of a ship with a robust total ship fiber-optic cable plant with limited network integration that is singularly managed and controlled by a distributed command function. Initial ships to consider a totally integrated design using an $(IC)^2$ architecture are: LPD 17, CVN 76, and the 21^{st} century combatant.

Introduction

The dawn of the Information Age, including the emerging "Information Highway" and the technologies associated with the integration, fusion, and distribution of heretofore unavailable data to the general population will radically change existing paradigms of ship design and acquisition. Affordable fiber-optic sensors, high-performance networks, fault-tolerant distributed computing, and advanced software tools that can develop, operate, and maintain highly complex systems with substantially fewer personnel will revolutionize the 21^{st} century Navy in a manner not unlike the transition from sail to steam. The conceptual definition of the future total ship system architecture and the engineering process that will realize the concept of the "Autonomic Ship"[1] are being heavily influenced by the new commercial off-the-shelf (COTS) products and de facto standards, particularly those based on extensive technology research and development occurring in the telecommunications industry. The $(IC)^2$ engineering validation, also known as "Wallops '93," conducted during September through

October 1993 at the Naval Surface Warfare Center's Combat System Engineering Facility (CSEF) located at Wallops Island, Virginia (Figure 1), demonstrated progress toward the goal of total ship integration using a commercially based open-system architecture. Wallops '93 also provided a flexible, adaptable, and cost-effective approach to developing key elements of an evolving total ship warfighting architecture.

Background

For the past 30 years, the design and acquisition process that implements change aboard naval surface combatants was based on the premise of an evolutionary, increasingly dangerous open-ocean threat. Military leadership was largely responsible for the development of the latest technologies and their implementation in diverse subsystem procurements, resulting in highly specialized warfighting capabilities for our ships. When technological change was generational, the long-term "vision driven" development approach for ships with a nominal 30-year lifetime was usually adequate to meet the fairly predictable threat. The decade of the 1990s has already witnessed the collapse of the Soviet Union and communism, Desert Shield/Storm, the civil war in Bosnia-Herzegovina, a series of crises in Africa, and brewing hot spots in Asia and the Caribbean.

The Navy/Marine Corps "From the Sea"[2] and the Joint Chiefs of Staff "C^4I for the Warrior"[3] strategies recognize the need to balance reduced joint forces with increasingly flexible options to deploy them on combat. Implementing such strategies requires a laboratory focus toward solving the engineering challenges posed by information fusion; tactical decision-making under stressful, quick-response situations; and development of robust communications methodologies to support adaptable, total ship functions/missions. The adoption of an extensible ship architecture, which creates an environment for affordable, rapid insertion of new capabilities as dictated by "come as you are" operational requirements, should become a priority.

2010 Surface Combatant—"H" Architecture

In 1988, the Report of the Ship Operational Characteristics Study on the Operational Characteristics of the Surface Combatant of the Year 2010 cited the "information intensive" nature of this future ship, detailing requirements for information management approaches to tactical, technical, maintenance, and administrative data; reduced

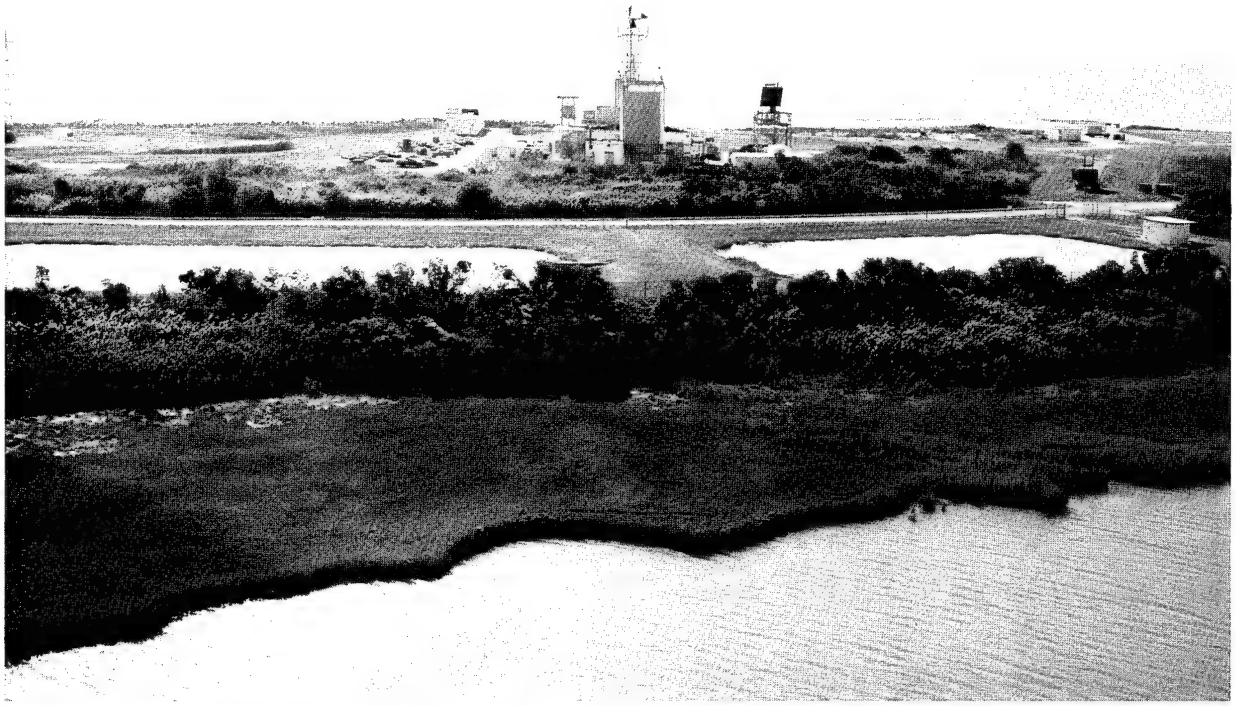


Figure 1. *Naval Surface Warfare Center Combat System Engineering Facility at Wallops Island, VA.*

manning; and affordability.[4] In 1990, VADM Jerry O. Tuttle, USN (Ret.), described a shipboard equivalent of the emerging Copernicus Navy command, control, communications, computers, and intelligence (C^4I) force architecture linked by the Tactical Data Information Exchange System (TADIXS) to a Communications Support System (CSS):

> "...Architectural vision for unifying ships' information needs in terms of both connectivity and data exchange to provide a logical framework for integrating ships as warfighting entities."[5]

Concurrently, Naval Surface Warfare Center Dahlgren Division (NSWCDD) completed a long-term analysis of existing and planned combat systems, including battle organization, projections of future surface warfare needs and technological opportunities, and emerging engineering requirements to partition increasingly large, complex systems. This effort resulted in the development of a combat system vision architecture (Figure 2, also known as the "H" Architecture).[6] The "H" Architecture included a dedicated system information coordination capability, which included functions not unlike those performed by the Space and Electronic Warfare Coordinator (SEWC), with a fully capable CSS in the future Copernicus Architecture:

> "The System Information Coordination (SIC) element, under battle conditions, interacts with Sensing and Warfare Coordination elements via the left-hand vertical Information paths to provide the Force and Ship Command elements with a summary of the tactical situation and environment. The SIC also controls sensing resources under Command direction when Warfare Coordination is not required. This role, under nonbattle conditions, extends to control of all sensing resources. In operational terms, this equates to performing the principal watch functions when the ship is not at battle stations."[6]

In mid-1991, the Navy approved a Non-Acquisition Program Definition Document (NAPDD #255-03-91) for $(IC)^2$ Development, which initiated an effort to coordinate the design of a total ship interior communications architecture for future naval surface ships. The scope included developing, integrating, simulating, and building prototypes; as well as testing system network hardware, firmware, and software products to complete and demonstrate a total $(IC)^2$ system emphasizing combat systems, logistics, and navigation for new ship designs to award starting in FY 95. System-level requirements included: (1) improving connectivity and interoperability, (2) streamlining network management, and (3) developing standard interface unit (SIU) hardware and software for voice and data applications. Figure 3 is a high-level depiction of the vision $(IC)^2$ total ship architecture, which considers the integration of the Copernicus and combat system vision architectures.

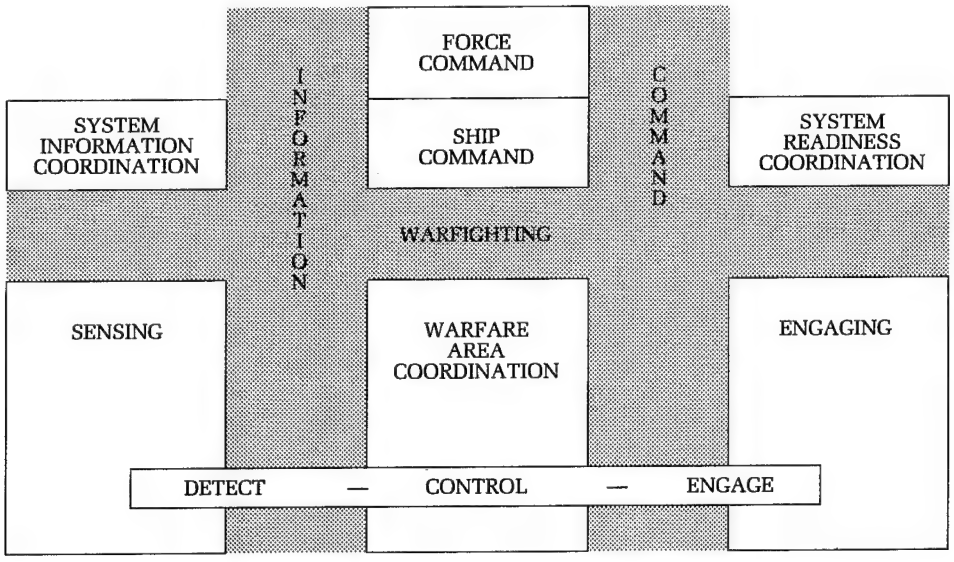


Figure 2. *Combat system concept.*

$(IC)^2$ Proof-of-Concept Validation

NSWCDD proposed an $(IC)^2$ proof-of-concept demonstration (also known as "Wallops '91") at the Wallops Island CSEF, based on the innovative development and early at-sea testing of the advanced development model of the Rapid Antiship Missile Integrated Defense System (RAIDS) aboard *USS John Hancock* (DD 981) in late 1990. Ultimately, the Naval Sea Systems Command (NAVSEA) selected

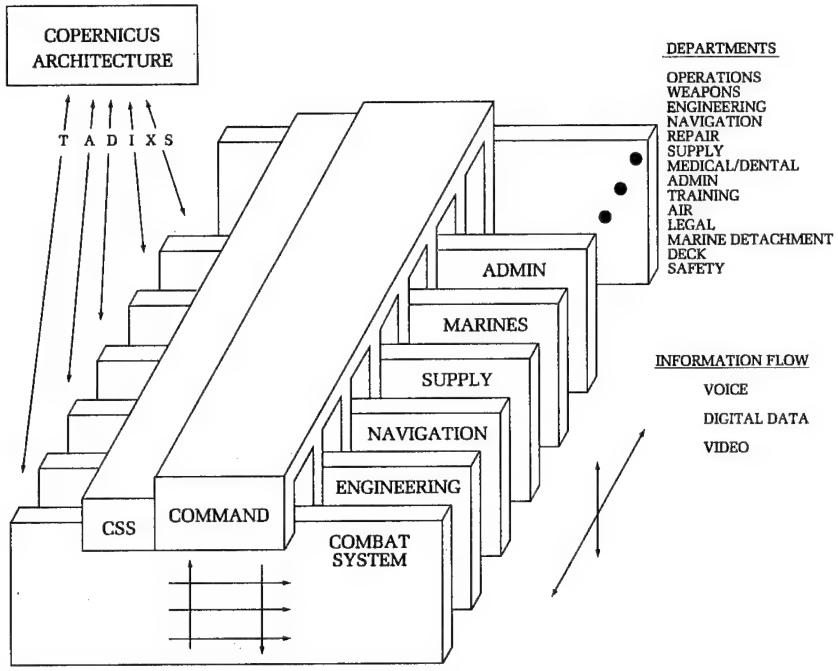


Figure 3. *(IC)² total ship architecture.*

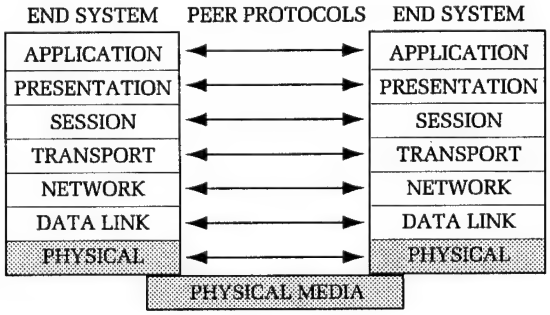
six individual network-based systems (Figure 4) to validate in Wallops '91 the concept of voice and data networks integrated into a single total-ship (IC)² system—a "network of networks."

As the system integration agent, NSWCDD designed and installed the fiber optic cable plant and directed the analysis of data transfer rates, message formats, and network protocols. As each of the networks was unique, working groups specified the hardware as well as the message protocols and formats for interconnecting networks. Ethernet was specified as the internet hardware because of its availability and low cost. While a higher bandwidth fiber-distributed data interface (FDDI) network (100 Mbps) was integrated, it was not a hardware driver. The well-known transmission control protocol/internet protocol (TCP/IP), based on the international standard open systems interconnection (OSI) reference model (ISO 7498), was selected and specified for all network traffic. Table 1 identifies the hierarchy of data communications protocols that constitute the OSI reference model.[7] Pseudo (not sanctioned by the Network Information Center (NIC)) Class A internet addresses were used by the integrated systems. The documented RAIDS message format was selected as a convenient basis for defining a limited set of internet broadcast communications using the transport layer user datagram protocol (UDP).[8] Figure 5 shows the Wallops '91 (IC)² fiber-optic cable plant.

In September 1991, using live aircraft services and Link 11 with the *USS Dwight D. Eisenhower* (CVN 69) Battle Group providing a realistic tactical scenario, the (IC)² system integration effort, lasting less than seven months from presentation of the initial RAIDS-based concept, successfully accomplished the following:

- Validation of the concept of separate voice and data networks integrated into a single fiber-optic networked system
- Concurrent use of FDDI, IEEE 802.5 token ring, and ethernet protocols; SAFENET nodes; and an integrated-services digital network (ISDN)-like network
- Simultaneous display on a single workstation of live AN/SPS-48C air search radar tracks; navigation data from the IBS, PLRS; and RAIDS engagement information
- Integrated Navy and Marine Corps tactical command and control data during a simulated amphibious assault
- Employing an open (i.e., nonproprietary) architecture, diverse systems were connected and integrated with little or no changes to system software
- A ship-to-shore link was exercised to demonstrate logistic and administrative

Table 1. *OSI Reference Model*



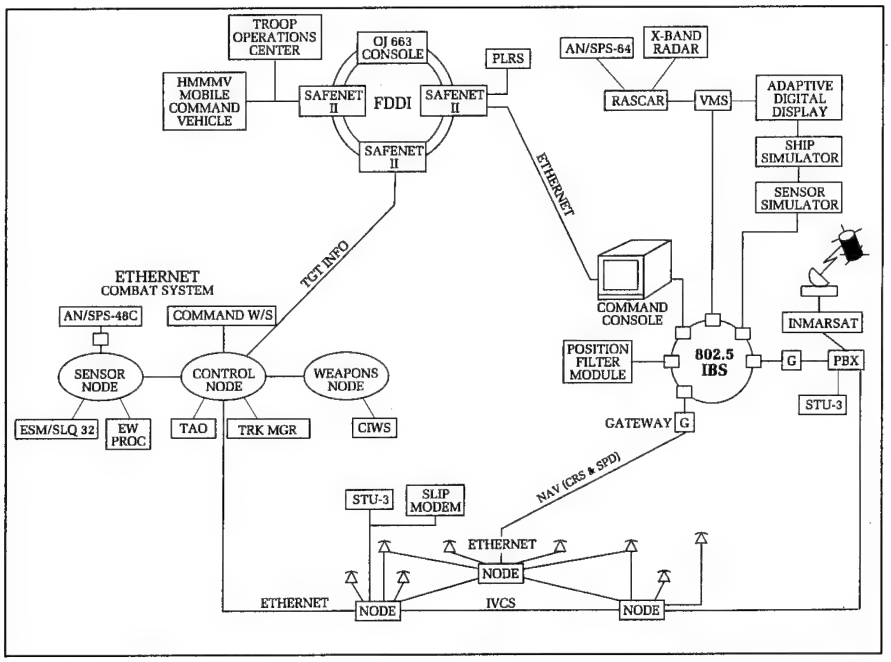


Figure 4. *(IC)² proof-of-concept composite.*[8]

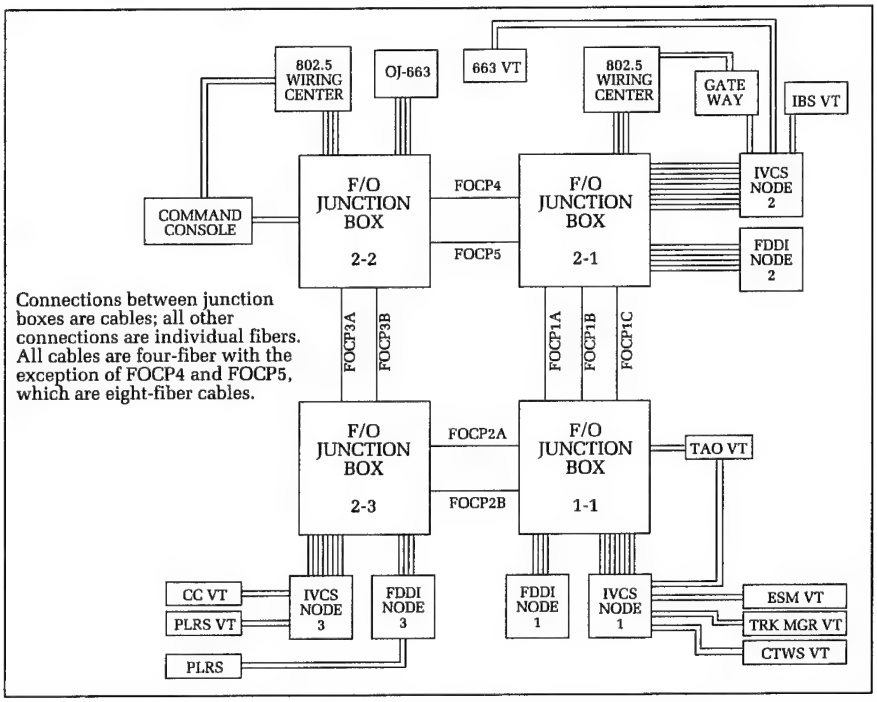


Figure 5. *(IC)² proof-of-concept cable plant.*[8]

data transfer to the defense data network via the INMARSAT satellite.

Despite these successes, it was evident that a more rigorous system engineering process would be required in order to develop the total ship (IC)² architecture.

(IC)² System Engineering

Wallops '91 encouraged NAVSEA to make a larger investment in a land-based engineering facility to continue the development of the (IC)² architecture. Accordingly, NSWCDD was directed to make modifications to the Wallops Island CSEF (also known as Building Z-41). During 1992, a two-story insert was constructed in Building Z-41's high bay, previously used by NASA for the assembly of medium-sized rockets. In early 1993, the installation of an extensive fiber-optic cable plant was begun to support the planned FY 93 validation testing. Using emerging fiber-optic standards for ship construction, an NSWCDD team of engineers created a robust, survivable mesh of fiber optic cables that could be linked by COTS passive and active routing components. By late spring 1993, the cabling of the entire Building Z-41 complex,

including potential locations for outdoor mobile laboratory elements, had been completed.

Fiber Optics

The $(IC)^2$ vision architecture prepares the ship for total connectivity, in part through a survivable, high-capacity, extensible fiber-optic plant. A summary of the advantages of fiber over conventional electric cables for data transmissions is useful in understanding the important role of a pervasive total ship cable plant in developing a future ship architecture:

- Higher transmission speed
- Increased transmission capacity
- Electromagnetic Environment (EME) Resistance
- Elimination of cable interference
- Increased cable runs
- Reduced bit error rates
- Reduced weight
- Simplified installation
- Improved security
- Data capacity growth
- Cost and supply
- Increased life
- Environmental resistance[9]

Combined with the development of future high-performance photonic switching fabrics, the potential bandwidth of optical fiber will be realized during the next century. The development of optical, multiple-access networks with architectures similar to current electrical frequency division multiple access (FDMA) systems, will provide an optical carrier operating in the hundreds of terahertz instead of the megahertz or gigahertz range of the electrical carrier.[10]

As the total ship fiber-optic cable plant design was evolving, and subsequently implemented at Wallops Island, an intensive effort was underway to define the $(IC)^2$ networking environment. The Naval Command, Control and Ocean Surveillance Center focused on the investigations associated with commercial routers, routing protocols, and a new area, network management tools, which had not been a key element of Wallops '91. NSWCDD launched four major technical focus areas: (1) development of an $(IC)^2$ standard message, (2) development of a network timing protocol, (3) analysis of appropriate network protocols, and (4) development of a standard network module. Each of these ongoing efforts played important roles in the evolution of the $(IC)^2$ total ship architecture. In particular, NSWCDD's activities were integral to the system integration of the "$(IC)^2$ Coalition" that would follow in FY 93.

Network Management

Basic or "Horizontal" network management was used during Wallops '93 on the Level 1 command network. This kind of capability is prevalent with today's systems; e.g., the typical administrative network manages primarily at the physical or, at most, the data-link level. These systems generally detect "hard" problems, which may actually cause a failure in the network. Nevertheless, "soft" errors, which do not cause system failures in an individual Level 2 network, cannot be detected but may accumulate and cause a general Level 1 problem. More extensive "vertical" management, attempted in only a few computer systems to date, attempts to look at the higher ISO layers as well but normally falls well short of identifying and resolving system problems for the operator. Still, this is an important next step in evolving network management. The vision total ship architecture should include a more integrated, robust capability that truly manages, not simply monitors:

> "...a network management system to meet end users' needs is that of an 'integrated' system. The 'integrated' world combines the 'horizontal' with the 'vertical.' This world will manage all subnetworks at all ISO levels...the capability will exist to handle the issue of soft errors...for a fully integrated view."[11]

Standard Messages

The development of the $(IC)^2$ standard message, also called the prototype (or "P") message, was based on fundamental data communication principles associated with implementation of an open, nonproprietary, network environment that could ultimately fuse as well as transfer data. Accordingly, the following objectives were established:

- Wide applicability
- Message content extensibility
- Reduced processing time for network user data extraction
- Data delivered in the format of the end users
- Data transferred in a format that optimizes bandwidth and data conversion

With these core design objectives in mind, NSWCDD engineers, working with their counterparts at other Navy activities, began an arduous 5-month process of defining data elements; e.g., seawater temperature, number of words, navigation status, etc., which were consolidated over another 3-month period by ship departments into a data-element dictionary. During this effort, standard measurement units for various entities (e.g., temperature, pressure, speed, etc.) were selected. The Naval Warfare Tactical Data Base and the existing set of well-documented interface design specifications for *USS Arleigh Burke's* (DDG 51's) combat system were used as references in this work. The department organization was chosen because of its logical relationship to the normal operation of all Navy ships. It was recognized that a capability was required to transfer important data during tactical situations and special evolutions that also had to be adaptable to the ship's battle organization and watchbills.

The development of an $(IC)^2$ baseline communications matrix, which defined generic relationships among command and the several ship departments, including an embarked department (e.g., Marine Corps units), was used as a vehicle to identify potential communication linkages (voice, data, video, analog, paper) that exist in a total ship $(IC)^2$ system. It was understood that these communications paths would change based on ship conditions of readiness, special operations, or during casualties. A top-level functional interface requirements document was produced from the communications matrix and was used to guide the initial process of developing the $(IC)^2$ standard message.

Defining and refining the specific P messages resulted in 51 $(IC)^2$ standard messages. These messages comprise approximately 450 data elements which were used during the engineering validation to integrate systems from 5 major departments (combat systems, air, engineering, navigation, and operations). These P messages provided an initial step towards consolidating and removing unique characteristics inherent in existing shipboard point-to-point interfaces. For example, current militarized computers, such as the AN/UYK-43 and AN-UYK-44, do not use the standard byte format employed by most modern processors. Without special purpose devices, such as a programmable network interface unit, this significantly limits their growth. In looking toward a future open-system architecture, proprietary network environments were deliberately ignored. Some messages might contain more information than required by a particular user. However, the alternative to one user extracting fewer data elements for his application was for all users to receive more messages. Adherence to the latter philosophy is considered inconsistent with the goal of information sharing, which is the primary advantage of a networked environment. A long-term view toward reducing software-development costs for $(IC)^2$ network users and controlling bandwidth use up front in the design process were important factors.

The P message is not particularly unique. Its format is based on the commercial External Data Representation (XDR) standard for describing and encoding data. P messages are also similar in format to those described in the currently underutilized NATO STANAG 4222, which specifies ways of representing various forms of digital data in terms of standard units of measure and data parameters. P Messages are compatible with multiple compilers (e.g., C, C++, Ada, and Fortran) and digital data processors. The P message is also independent of protocols (e.g., FDDI, Express Transfer Protocol (XTP), and Asynchronous Transfer Mode (ATM)), lending itself to accommodate growth and change in future networks.[12]

Network Timing

The second major technical thrust involved experimentation to determine the best means of maintaining an accurate clock for potentially hundreds of users and millions of sensors on a future $(IC)^2$ network. This is not a trivial requirement, considering that operating a 21st century Autonomic Ship as a "warfighting entity" will require precise synchronization of ship functions in response to evolving threats, enabling the transition from various conditions of readiness while allowing a smaller crew to rapidly adjust to unusual ship mission configuration or casualty situations. An analysis of the existing Network Timing Protocol (NTP), developed by the University of Delaware under Advanced Research Projects Agency (ARPA) sponsorship, revealed two major problems: (1) the software required significant computer-processing resources to execute, and (2) there were no provisions for rapid resynchronization. Therefore, the technical tradeoff was made to significantly streamline the code while maintaining the important NTP algorithmic features. This is clearly an area that requires significantly more research and development. An approach that would have established a parallel point-to-point timing network was rejected in order to investigate the issues associated with time synchronization of multiple applications on diverse networks. Testing was conducted to measure the drift

rate of several computer clocks operating on an ethernet network while quantifying round-trip delays of various signals and comparing them to a system time maintained by a master clock on the network.

An advanced Silicon Graphics work station, which could provide measured system time signals to the nearest 1 millisecond, was used as a time server for the clients on the network. (Note: Most hardware systems can normally measure time, at best, to a 10 millisecond accuracy. Industry is moving closer to the 1-millisecond capability with special hardware that can be used by a number of workstations and personal computers. A stated goal by some companies in the timing community is 50-nanosecond accuracy of a future network.)

The results of the timing experiments, which were normally conducted over 24- to 72-hour time periods, revealed that a 1-minute periodic resynchronization of clocks on the network would guarantee an accuracy of ±1 millisecond approximately 96.3 percent of the time and ±2-millisecond accuracy 99.9 percent of the time. This was deemed adequate to provide timing services for the Wallops '93 engineering validation. As a result, the server/client network module was distributed to those activities for preliminary integration and testing before arrival at Wallops Island.[13]

Network Protocols

Selecting the network protocol was accomplished with relatively little analysis, recognizing that network protocols for distributed real-time systems were largely under development or could not be implemented in affordable commercial systems that were widely available. For example, the lack of a product that provided guaranteed multicast communications, a key requirement for future networked combat systems, was a factor in the selection process. Again, TCP/IP was chosen for its low cost and availability. Nevertheless, the perceived requirement for new, improved protocols was tempered by research that indicated that the complexity of the TCP/IP protocol is not the most significant factor in processing delays for systems operating on a network. In fact, some experts contend that application system bottlenecks can be minimized and throughput significantly improved through fixes to other key factors such as "...I/O service, interrupts, task switching, and scheduling...."[14] The use of UDP was not permitted. This meant that only connection-oriented, end-to-end, reliable byte transmission would be authorized in the Wallops '93 $(IC)^2$ network. In addition, the Level 1 and Level 2 networks were each assigned a block of Class C internet addresses officially sanctioned by the NIC.

Application Programmers Interface

The most significant technical accomplishment during the system development phase was the conceptualization and creation of a network module, which has since evolved to an applied patent for a network Application Programmers Interface (API). The purpose of the network API is to provide a potential user of the $(IC)^2$ total ship network with the requisite tools to be able to access the network efficiently, make connections with other systems on the network, maintain communications, handle error conditions, process the $(IC)^2$ standard messages, and provide message encapsulation services for those systems that may not need to use a standard network messaging scheme but desire to be in a fiber optic networked environment. More importantly, the system developer does not have to be knowledgeable in network communications and can focus on optimizing his application program for the required ship function(s).[15] In proprietary networks, this capability is closely held, and changes to the network are costly. For so-called "stovepiped" systems (this is the general situation aboard today's ships), the system developer does not usually make design decisions based upon use of his network by other applications. Therefore, the network APIs are tightly coupled to the application program, normally requiring significant software changes to integrate with other systems:

> "Although distributing services throughout networks of autonomous computers offers many potential benefits, developing distributed applications is more complex compared with standalone applications. A significant portion of this complexity arises from limitations with conventional tools and techniques used to develop distributed application software. In particular, many standard UNIX networking mechanisms (such as sockets and the Transport Layer Interface (TLI)) and reusable component libraries (such as ONCRPC) lack type-safe, portable, re-entrant, and extensible application programming interfaces."[16]

Figures 6 and 7 show the network API and the relationship of P messages to the network

applications. The network API code was provided to all participants to link into their application programs. As a result, most Level 3 systems were able to be integrated on the $(IC)^2$ network in a matter of hours or days. For those systems that developed their own network interfaces and elected not to use the network API, integration periods of several months were not unusual.

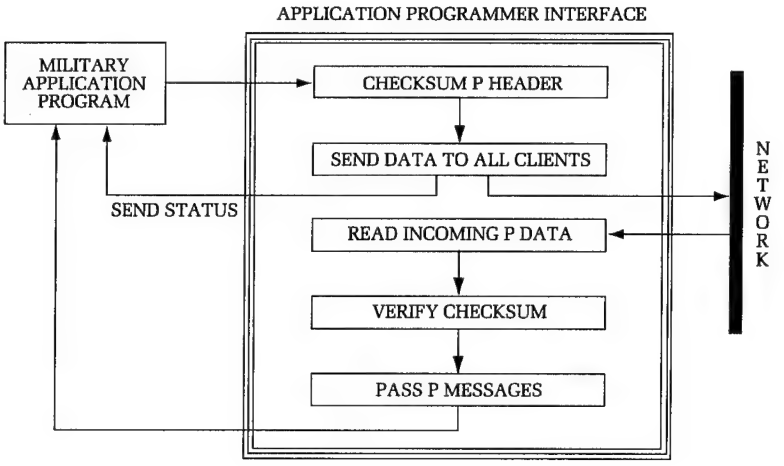


Figure 6. *Network API.*

$(IC)^2$ Engineering Validation

In the early summer of 1993, the initial $(IC)^2$ architecture engineering validation preparations were completed. System integration utilizing techniques described above were implemented by the growing "$(IC)^2$ Coalition" and continued at Wallops Island.

The system configuration for Wallops '93 consisted of a command (Level 1) FDDI network and six department (Level 2) networks of various technologies (FDDI, ethernet, and ISDN). A parallel Level 1 voice ISDN network was also utilized. Satellite and landline T-1 connectivity to various government and contractor facilities located throughout the country was also used to demonstrate the relative ease of system integration using $(IC)^2$ methodologies. The final configuration for the engineering validation is depicted in Figure 8, a total ship system of approximately 44 subsystems provided by 26 Navy activities and contractors. A realistic amphibious operational scenario was devised using both live and simulated sensor inputs that exercised the $(IC)^2$ architecture as it would be employed on a DDG, CV, or LX (now LPD-17) ship through retrofit, upgrade, or new development.

An innovative Alternate Path FDDI (APFDDI) network based on research conducted at NSWCDD was also demonstrated. This network incorporated enhanced survivability features for future networks using commercial concentrators and available FDDI protocols. (Note: In conventional FDDI networks, two faults will isolate portions of the net. Using APFDDI, an FDDI network can survive up to three errors before isolation of sections of the network occur.)[17] Network video teleconferencing based on techniques developed by the Internet community on the worldwide experimental Multicast Backbone

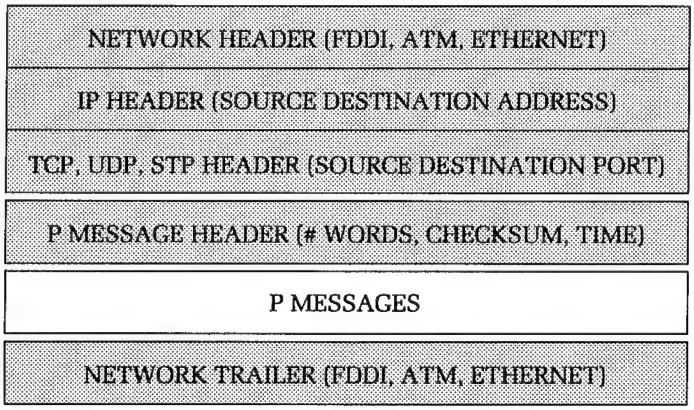


Figure 7. *Encapsulation of P messages.*

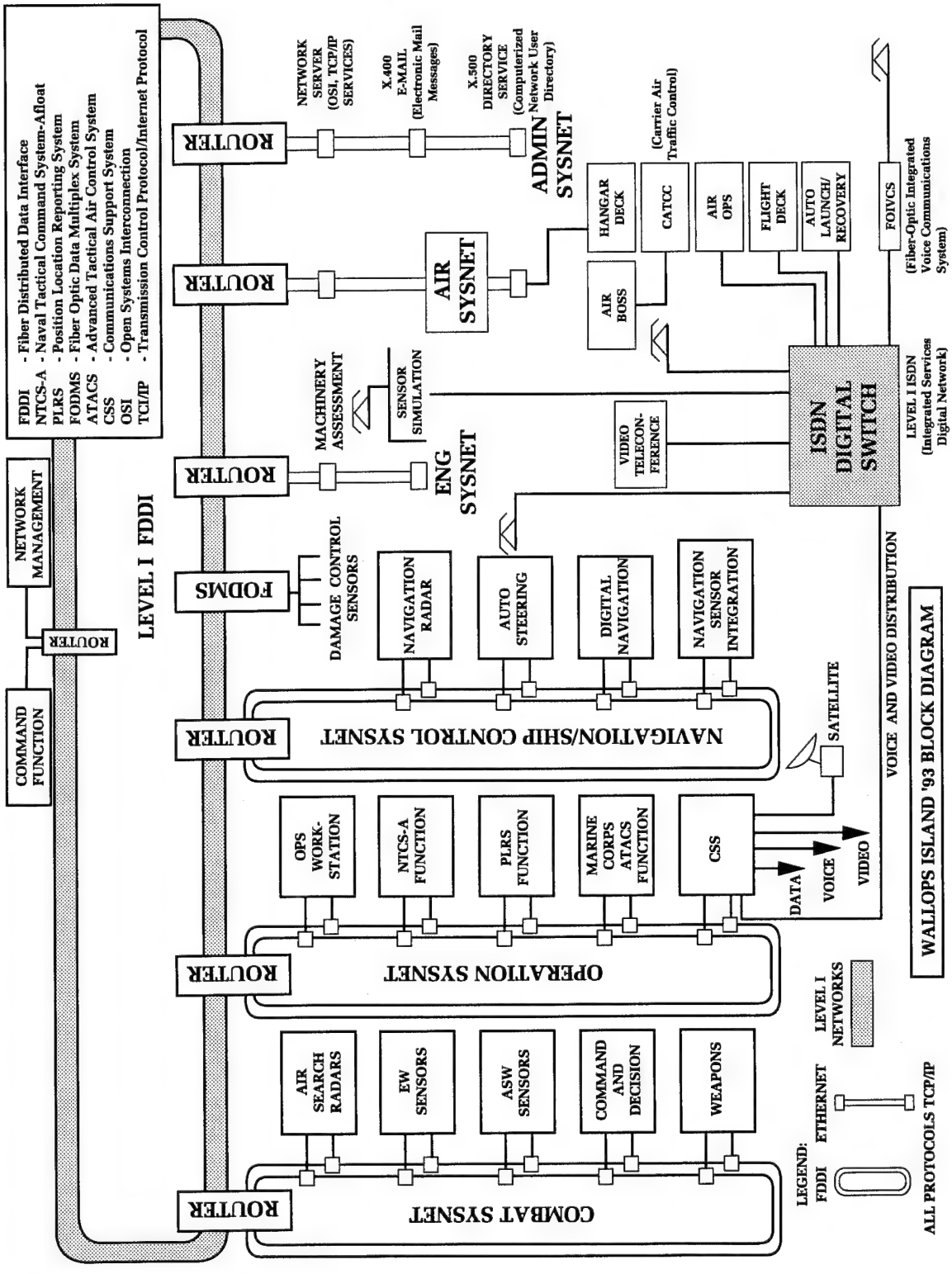


Figure 8. *Validation architecture.*

network (also called the "MBONE") was demonstrated. Worldwide mail messaging using the emerging X.400 and X.500 international standards was also provided. Additional experiments with the actual NTP were conducted using hardware that provided 1-microsecond precision for selected workstations.

Engineering Tests

Connectivity: The Level 1 connectivity tests verified the integrity and continuity of the FDDI Level 1 network fiber by using light sources and power meters. The tests verified that the signal path was continuous with no breaks or flaws that would impair transmission of LAN traffic. It proved that the established fiber path through routers, trunk coupling units, path panels, and other devices on the path did not prohibit or impair message flow. Level 2/3 connectivity tests verified the connectivity of the Level 2 and 3 systems on the Level 2 networks and connectivity with Level 1. These tests verified continuity of signal paths and interoperability of selected components by sending and receiving P messages, alerts, or E-mail.

Latency: The technique used to measure latency involved sending a time-stamped message from one work station to a second work station, which immediately returned the message and, subsequently, calculated the round-trip delay time. Numerous tests were conducted from various locations on the $(IC)^2$ network, but more substantive analysis and testing of real-time systems is required. If some reasonable assumptions are made, the round-trip delay time can be considered to be twice the latency. The calculated average network latency, as determined from a series of tests, is shown in Figure 9.

Data Accuracy: The ability to maintain information integrity as messages are sent from one station to another across the network—data accuracy—was examined by two methods. The first method was the observation of displayed information on the various workstations as the systems were integrated and during conduct of demonstrations. The second method was the analysis of collected data to examine input and output error rates. The first method looked at the accuracy of the data itself, while the second method looked at the network architecture's ability to transfer the data accurately. Data accuracy was examined by comparing received data with transmitted data. This procedure was performed throughout integration of the various Level 3 systems within the $(IC)^2$ architecture. Data collection for later analysis of the architecture's ability to accurately distribute data was conducted in conjunction with running the demonstration scenario. The output packet error rates were extremely low, with the highest rate found being well below one-hundredth of one percent.

Environmental problems: Electromagnetic interference from the AN/SPS-48C radar adjacent to Building Z-41 initially caused significant operational problems with unshielded COTS equipment, such as the Level 1 routers. Installation of shielding on the building minimized this problem, but many participants continued to point to EMI as a source of interference to communications. This should not be a significant problem aboard ship if proper selection of equipment location is combined with proper EMI prevention techniques.

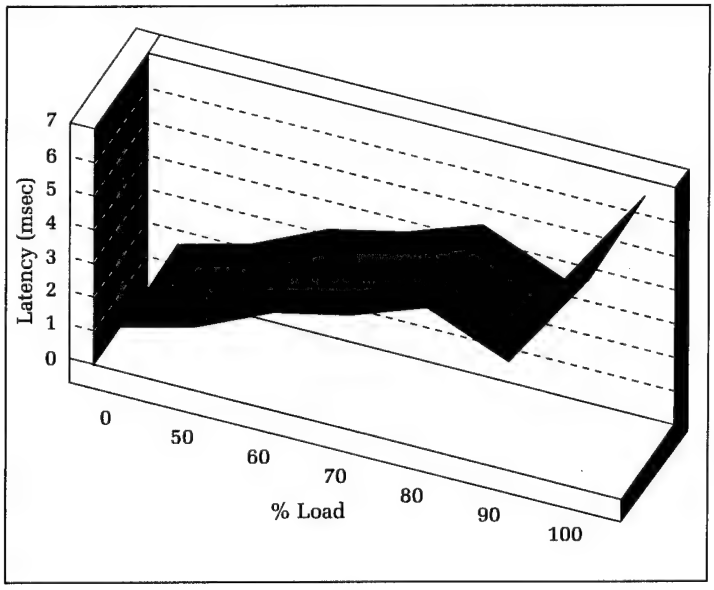


Figure 9. *Calculated average latency.*

Another problem was indicated by test data. The data indicates that some device on the network was transmitting bad data or not recognizing protocol messages that were transferred across the backbone. As the site was dismantled, the source was never identified. This points to the need for a more robust network management, analysis, and readiness capability than was provided with the commercial products tested at Wallops Island.

Network Load Limits: Architecture capacity was examined by deriving Input Data Rate (IDR) and Output Data Rate (ODR) information from collected data. Architecture capacity was examined by comparing router traffic in bits per second with design capacity of FDDI (100 megabits/second) or design capacity of Ethernet (10 megabits/second). This procedure was performed numerous times under various conditions. Typical loading of Ethernet rarely exceeded 1-percent utilization. FDDI loading on any interface rarely reached 0.05-percent utilization. FDDI backbone utilization (the sum of all IDRs and ODRs on the FDDI as measured on one of the ports of each of the Level 1 routers) averaged approximately 1-percent utilization. Figure 10 demonstrates available versus used capacity, clearly demonstrating the potential growth of an advanced $(IC)^2$ Total Ship Network.[18]

Information Fusion

Information fusion was a key demonstrated capability of the $(IC)^2$ engineering validation. This was accomplished in an unclassified network environment by a combination of standardization of the information elements and fusion of the various data sets through new correlation algorithms and summary display techniques. For example, the Global Correlation Engine (GCE), the product of extensive 6.2 Research and Development initiatives, was integrated on the operations system network and provided a fused tactical picture to all stations requiring this information. Using an innovative combination of a Gaussian correlator integrated with a non-Gaussian tracker and multiple hypothesis correlator-tracker, the GCE built a complete tactical picture based on all available organic, nonorganic, real-time, nonreal-time, local, and over-the-horizon track information.[19] Summary screens of system status for the Level 2 networks were combined into a command summary screen on the Level 1 network. These techniques were also used to readily provide additional functions to the Command Station, which is being deployed on a number of surface combatants. Some workstations used voice, data, and video information to provide state-of-the-art virtual reality presentations, touch-screen displays in an X-windows environment, and touch-panel functional controls.

Conclusions

The total ship architecture for the 21st century Navy is being framed by emerging key

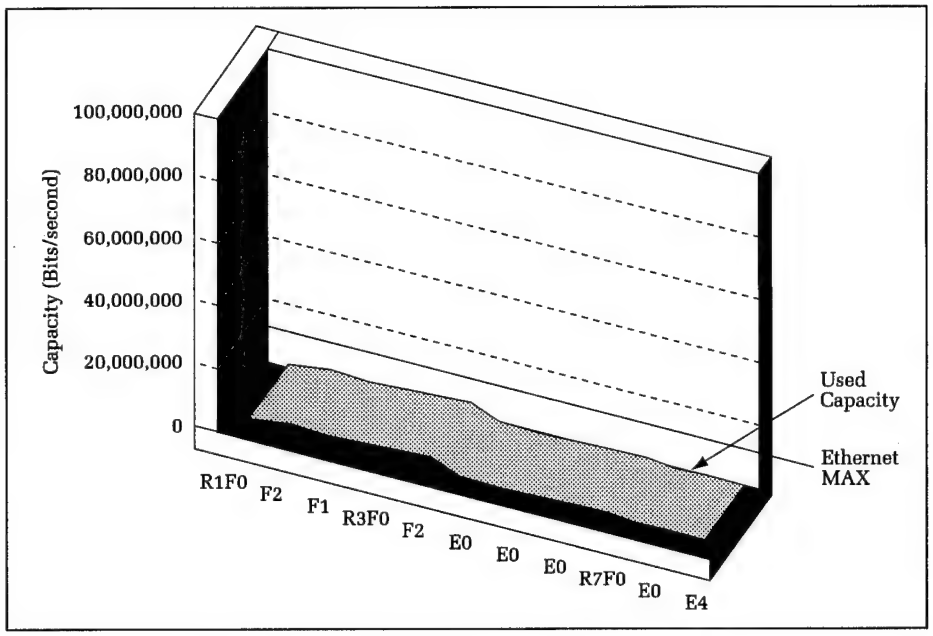


Figure 10. *Architecture remaining capacity.*

system integration technologies in the telecommunications industry. $(IC)^2$ has provided an initial concept, architecture, and system engineering approach that can be critically reviewed and affordably evolved to a total ship engineering environment. The $(IC)^2$ architecture engineering validation conducted at the Naval Surface Warfare Center's Wallops Island CSEF and distributed sites throughout the country integrated ship functional departments such as combat systems, operations, navigation, engineering, air, administration, and command. This was accomplished by using a COTS hardware and software environment that readily supported the standardization and fusion of ship information (e.g., data, voice, mail, and video).

Wallops '93 included demonstrations that were conducted for over 500 Flag officers, senior executives, and other visitors from 4 through 22 October 1993. The outcome of the Wallops Island '93 effort was an evolution of the $(IC)^2$ architecture to a stage that is mature enough for design and construction of a ship with a robust total ship fiber-optic cable plant with limited integration that is singularly managed and controlled by a distributed command function. Evolution of the architecture in a "build a little, test a little" prototyping approach will provide a basis for affordable integration of information technology as it matures. Leading candidates for initial installation of the $(IC)^2$ architecture are: LPD 17, CVN 76, and the 21st century combatant.

The $(IC)^2$ architecture validation was conducted within an unclassified environment with multilevel network security issues deferred for later evaluation. Integrated network management for survivable total ship mission performance and reconfiguration in response to real-time tactical situations remains an important research area. Critical linkages between the combat system and hull, mechanical, and engineering communities to operate the ship as a "Total Warfighting Entity" need to be seriously system engineered together using available and near-term telecommunications products in fiber optics, high-speed switching, and routing, with emerging advanced distributed computing and software development tools. These critical experiments should be conducted with an objective to revolutionize—and prototype—the processes by which we build and acquire ships in the next century.

References

1. Smith, T. D., Ditizio, F., Hoyle, S. and Wilson, R. A., "Autonomic Ship Workshop Results," *Carderock Division, Naval Surface Warfare Center Innovation Center Technical Report* (CRDKNSWC-INCEN-94/001), Nov 1993, p. 9.
2. O'Keefe, Sean, Kelso, Frank B., and Mundy, C. E.; *...From the Sea, Preparing the Naval Service for the 21st Century*, Sep 1992.
3. Edmonds, Albert J., *C4I for the Warrior*, 12 Jun 1993.
4. CNO, OP-03K; *Report of the Ship Operational Characteristics Study on the Operational Characteristics of the Surface Combatant of the Year 2010*, (Volume I), Serial 03/8S577900, 26 Apr 1988.
5. CNO, OP 94; *Navy-wide Data and Network Protocol Standards*, CNO ltr 3093 Ser 940D/OU540322, 6 Jun 1990.
6. Cullen, Richard P., *A Concept for Future Force Combat Systems: An Enduring Structure*, Naval Surface Warfare Center, Dahlgren, VA, NAVSWC TR 90-121, Feb 1990.
7. Bertine, H. V, "Physical Interfaces and Protocols," *Computer Network Architectures and Protocols*, C. Sunshine (Editor), Plenum Press, New York, 1989, p. 40.
8. Carle, G. L. and Irey, P. M, *$(IC)^2$ Concept Validation Final Report*, Naval Surface Warfare Center Dahlgren, VA, NSWCDD/TR-92/69, May 1992.
9. Freer, John, *Computer Communications and Networks*, Pitman Publishing, London, 1988, pp. 31-33.
10. Hinton, H. Scott, *An Introduction to Photonic Switching Fabrics*, Plenum Press, New York, 1993, p. 128.
11. Skrzypczak, Casimir S., "Network Management Alternatives for the Future," *Network Management and Control*, Edited by A. Kershenbaum et al., Plenum Press, New York, 1990, pp. 4-5.
12. Ratway, Michael J., "Interface Design Document for the Integrated Interior Communications and Control 1993 Validation Communications Messages," *Dahlgren Division, Naval Surface Warfare Center Report* ((IC)2-IDD-000), 4 Nov 1993, pp. A1-A7.
13. Do, Tuy T., *Network Time Synchronization Experiment*, Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA, NSWCDD/TR-94/177, Jul 1994.
14. Kanakia, Hemant, "Transport Protocols for High Speed Networks," *Recent Technical Developments in Telecommunications*, P. Levaux (Editor), Elsevier Science Publishers B.V., Amsterdam, 1992, pp. 118-119.
15. Plakosh, Daniel, "Network Application Programmers Interface," Navy Patent Application 76232, 11 Apr 1994, p. 2.
16. Schmidt, Douglas C., "The Adaptive Communication Environment: Object-Oriented Network Programming Components for Developing Distributed Applications," in *11th and 12th Sun User Group Conferences*, San Jose (7-9 Dec 1993) and San Francisco (14-17 Jun 1993).

17. Hiles, William S. and Marlow, David T., "Experimentation on the Concentrator Tree with Loopback," *Proceedings of the 18th Conference on Local Computer Networks*, Minneapolis, 19-22 Sep 1993, pp. 147-156.
18. Mangleburg, C. M., *$(IC)^2$ Architecture Validation at Wallops Island*, Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA, 28 Mar 1994.
19. Lynch, James P. and Monach, W. Reynolds, "Non-Gaussian Fusion in a Stressed Network Environment," 1994 Command and Control Symposium, Monterey, CA, 21-23 Jun 1994.

The Author

FRANKLIN A. ERVIN graduated from the United States Naval Academy in 1975 with a B.S. in Ocean Engineering and was commissioned as an Ensign in the U.S. Navy. He served on active duty for 8 years in a variety of at-sea as-signments in surface ships before coming to NSWCDD in 1983. He is a graduate of the Navy's Nuclear Propulsion School and Surface Warfare Officer School, Head of Department Course. He obtained a Masters Degree in Public Administration from the Virginia Polytechnic Institute and State University in 1986 and graduated from the U.S. Naval War College in 1987. He is also a Commander in the U.S. Naval Reserve. He is the Head, Combat Systems Engineering Branch in the Ship Defense Systems Department and also serves as the Program Manager for the relatively new research and development initiative, the Integrated Interior Communications and Control Program. His primary research focus is in developing and prototyping telecommunications and information fusion technologies, which can integrate a broad spectrum of functional requirements in a Navy ship environment.

Tactical Ballistic Missiles Trajectory State and Error Covariance Propagation

Gregory H. Drescher and Gary L. Sitzman

Tactical Ballistic Missiles (TBMs) are fast becoming a threat to world peace. To counter this threat requires the development of a new concept—one that combines the traditional players of target identification, communications, fire control, and the antiballistic missile. In this type of engagement, the total process is played on a compressed time line. Not minutes, but seconds, are the difference between success and failure. As a result, the time needed to search and detect an incoming TBM must be kept to a minimum. To help attain this goal, the TBM's time, position, velocity, and error covariance information can be passed over to the defending platform from a remote tracking platform via communications link. This information will aid the Anti-Tactical Ballistic Missile (ATBM) platform in constructing a fast and efficient search and detect strategy.

This article addresses tracking TBMs from space and surface radar platforms. Modeling errors in the tracking system produce an uncertainty covariance matrix. This article identifies the minimum number of covariance elements that are needed to produce an efficient search strategy. Given that the ATBM platform has received the TBM's state via communications link, a trajectory propagation model is needed to predict impact parameters and uncertainty error ellipses. This article quantifies the error introduced by various trajectory models and identifies the model fidelity required for a given set of TBMs.

Introduction

TBM error covariance and trajectory-state information are essential to any ATBM platform. The reason is simple—time is critical. With no information at all, an ATBM platform must, at the very least, search a defended sector of the sky. With a high-energy waveform, the ATBM's radar may reach hundreds of kilometers, but some TBMs may be incoming at a few kilometers per second at the extent of the radar range. This roughly equates to a few minutes to detect, intercept, and destroy the incoming TBM. While this is happening, the system must continue to search for more incoming TBMs. With a cue from another tracking platform that contains only nominal trajectory information (with errors present) about the incoming TBM, the ATBM platform can propagate this trajectory information forward to the extent of its radar. The ATBM's radar could then be directed at the nominal trajectory location, but if the TBM is not

detected at the nominal position, the search continues with no information about where to look next. A cue that contains trajectory information along with error covariance information will allow for the creation of a search strategy that dictates where the TBM is likely to be found and, more importantly, will dictate where the TBM will not be found.

Error covariance information can require a relatively large amount of space in a communications channel. This is compounded when many TBMs are in the air at one time; therefore, it is important that only the critical error information be sent. This study will address only a small part of the total problem of tactical ballistic missile defense (TBMD). Tracking accuracy, communication of accuracy data, and propagation of state and accuracy information will be discussed. A set of TBMs with ranges between 200 and 3000 kilometers is the threat of concern.

Tracking Tactical Ballistic Missiles

Satellite Tracking

Tracking from space with satellites can certainly be effective in identifying missiles in flight. Infrared (IR) sensors will detect missile plume. Continuous tracking is not necessary, but it is critical to track at a high rate during burnout. The largest error arises from the acceleration undetected between last track and burnout.

Tracking from satellites in space is depicted in Figure 1, where passive IR sensors are employed. The uncertainty in target state (position, velocity, and acceleration) is illustrated by error ellipsoids about the estimated mean state value. The target's estimated state and uncertainty will be sent to communication down links for reception by ATBM platforms.

A set of TBMs was used in this study. Acceleration profiles for each TBM were integrated from launch to burnout. Burnout conditions were then propagated to impact under sample gravity fields. A grid of launch points, several flyout azimuths, and satellite measurement parameters depicted in Figures 2 and 3 were used with each missile to determine the missile-tracker geometry effects.

Errors were propagated using state error equations that are shown at the top of Figure 4. The time derivative of the position error is equal to the velocity error. The time derivative of the velocity error is equal to the sum of two terms. One term is due to off-nominal gravity effects caused by position error. This was modeled by multiplying the position error by an inverse-square gravity gradient matrix, $\overline{(J)}$. The other term is just due to the unmodeled portion of gravity free acceleration, $\overline{f(a)}$. These equations were integrated to form a matrix of numerical partial

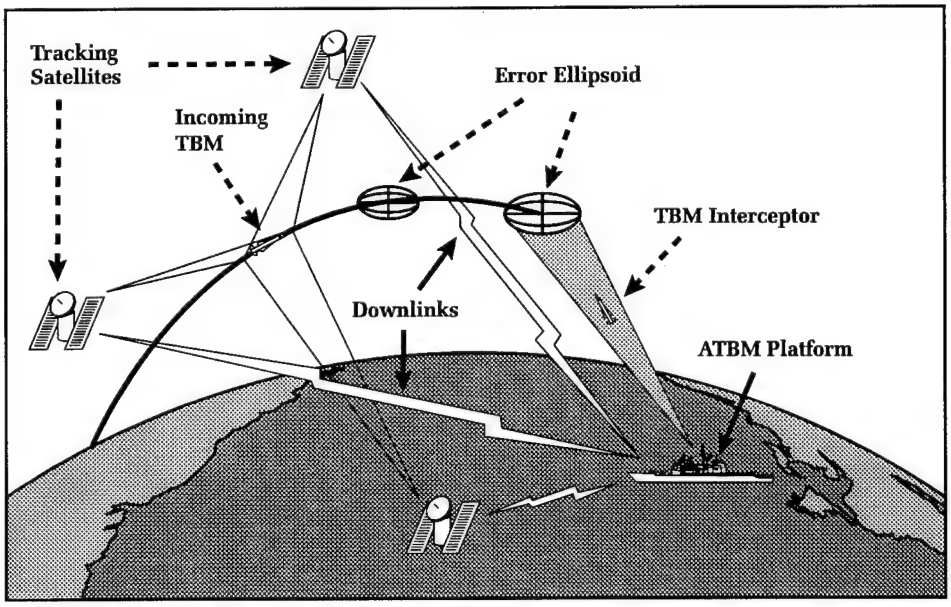


Figure 1. *Tactical ballistic missile defense.*

Satellite Measurement Sigma - σ_m of satellite measurement system (1000, 2000, and 5000 feet)

Launch Point - Grid (Latitude 0, 30, 60; Longitude 0, 30, 60)

Flyout Azimuth - 0, 45, and 90 degrees from north

Update Spacing - Last update time x seconds before burnout

Number of covariance elements to be transmitted at time of last update.
(Full 36 (6 by 6), 6, and 3)

Figure 2. *TBM trajectory parameters.*

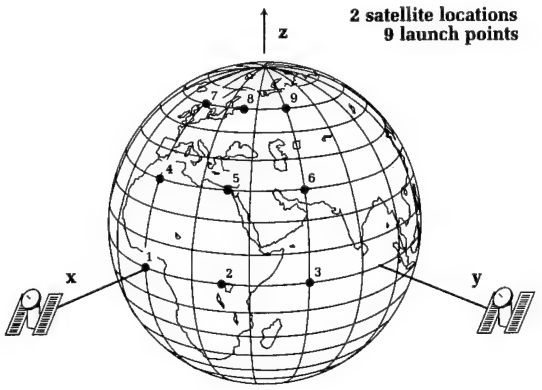


Figure 3. *Tracking and estimating state covariances of TBMs.*

derivatives that was used to "transition" the error covariance matrix from one trajectory time to the next (Figure 5). Method 1 transition included three-dimensional position, velocity, and acceleration scale factor errors, SF_x, SF_y, and SF_z, resulting in a 9 by 9 covariance matrix. The elements of the covariance matrix include variances and covariances of position error, velocity error, and scale factor error.

Error equations in method 1 assume no correlation among acceleration components ax, ay, az, and treat each component of acceleration as an independent error source. When the error equations of method 1 are used, the resulting error ellipsoids grow along the principal axes in proportion to the direction and magnitude of acceleration. The principal axes of the ellipsoids do not change their orientation until long after burnout when the gravity gradient term has had time to propagate.

Method 2 assumes correlation is known between acceleration components and considers acceleration magnitude as an independent error source. The single scale factor corresponding to acceleration magnitude error, contributes to position and velocity errors in the x, y, and z directions simultaneously. This causes the error ellipsoids to "elongate" along the direction of the thrust and results in a 7 by 7 covariance matrix.

The measurement update is performed instantaneously by a simple Kalman filter with no process noise (Figure 5). Other error sources include initial state uncertainty, unknown acceleration between last track and burnout, and tracker measurement error.

State error equations: $\dot{\overline{\Delta R}} = \overline{\Delta V} \qquad \dot{\overline{\Delta V}} = \bar{J} \bullet \overline{\Delta R} + \overline{f(a)}$

Method 1:

$$f(\bar{a}) = \begin{pmatrix} SF_x & 0 & 0 \\ 0 & SF_y & 0 \\ 0 & 0 & SF_z \end{pmatrix} \bullet \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

- 3 scale factors (x, y, and z) are evaluated as independent error sources

State Covariance:

$$\begin{pmatrix} \sigma_R^2 & \sigma_R\sigma_V & \sigma_R\sigma_{SF} \\ \sigma_V\sigma_R & \sigma_V^2 & \sigma_V\sigma_{SF} \\ \sigma_{SF}\sigma_R & \sigma_{SF}\sigma_V & \sigma_{SF}^2 \end{pmatrix}_{9x9}$$

Method 2:

$$f(\bar{a}) = (SF) \bullet \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

- Only 1 scale factor used as an error source (acceleration magnitude)

State Covariance:

$$\begin{pmatrix} \sigma_R^2 & \sigma_R\sigma_V & \sigma_R\sigma_{SF} \\ \sigma_V\sigma_R & \sigma_V^2 & \sigma_V\sigma_{SF} \\ \sigma_{SF}\sigma_R & \sigma_{SF}\sigma_V & \sigma_{SF}^2 \end{pmatrix}_{7x7}$$

Figure 4. *Error modeling for TBM state estimation.*

Kalman Gain Matrix:

$W = C\Theta^T(\Theta C\Theta^T+R)^{-1}$

C = System Covariance matrix
Θ = Measurement matrix
R = Tracker Covariance matrix
I = Identity matrix
W = Kalman Gain matrix
Φ = Transition matrix

Covariance Update:

$C_{(+)} = (I - W\Theta)\, C_{(-)}$

Covariance Transition:

$C_{(t1)} = \Phi_{(t0 \rightarrow t1)} C_{(t0)} \Phi^T_{(t0 \rightarrow t1)}$

Figure 5. *Standard Kalman filtering approach with no process noise.*

The IR tracking measurement was simulated as the target's position error, perpendicular to the tracker's line of sight (LOS) to the target (Figure 6).

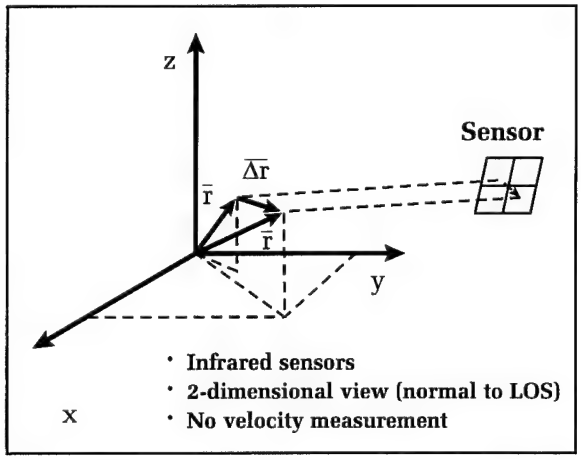


Figure 6. *Satellite sensor model.*

Results from space IR tracking are shown in Figures 7 through 10. Method 1, uncorrelated acceleration components, results in a mostly spherical error ellipse, with component size proportional to acceleration magnitude. Figure 7 shows root sum square (RSS) position sigma versus time for the nine-launch-point grid. Solid lines represent one-sigma errors resulting from propagating the full 81-element covariance. Dashed lines represent one-sigma errors resulting from propagating a covariance matrix in which all elements were zeroed out just after burnout, with the exception of the position and velocity variances. The position variances were set equal to the square of the satellite measurement sigma, and the velocity variances were left as is. This process simulates placing only the three velocity variances on the communication link instead of the full 81-element covariance matrix.

Figure 8 shows an X-Z plane cross section of two ellipsoids at reentry, generated using method 1 equations and the previously described process for transmitting only three velocity variances. The difference between the full covariance and the three-element covariance is too small to be seen at this scale.

Method 2, known correlation with unknown magnitude, results in error ellipses elongated along the acceleration vector. The covariance matrix generated by this method can still be regenerated using only three velocity variances, but the variances must be transmitted in a different reference frame.

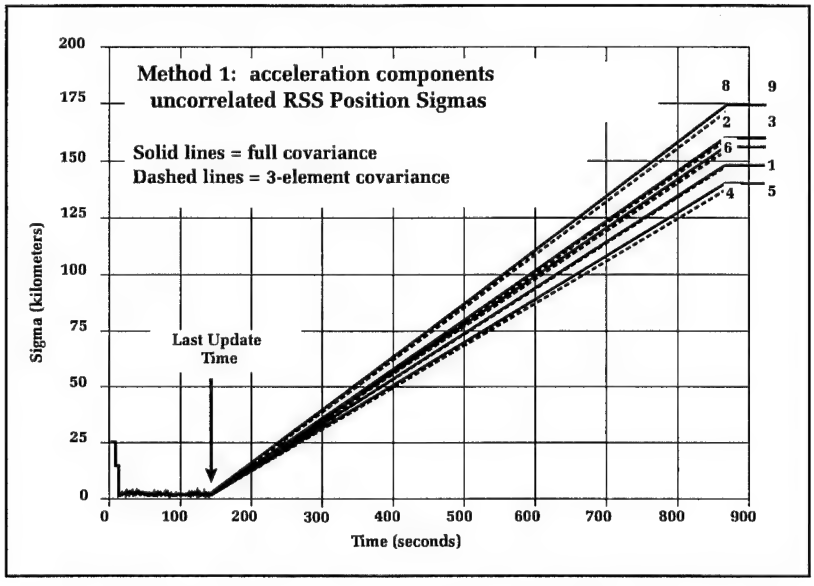


Figure 7. *Position uncertainty vs time for method 1 modeling.*

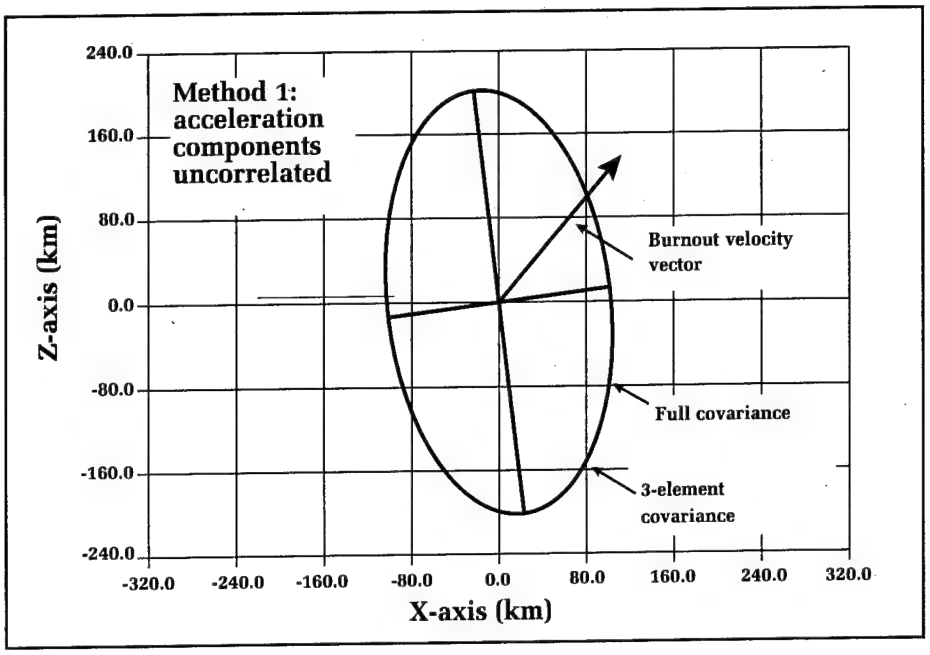


Figure 8. *Position ellipses at reentry, TBM-2 in X-Z plane, update interval 10 seconds, two satellites on X and Y axes, 13 simultaneous fixes, burnout at 138.5 seconds.*

Figures 9 and 10 show an X-Z plane cross section of ellipsoids at reentry, propagated using method 2 equations. The more rounded ellipse results from zeroing out the off-diagonal covariance elements as described earlier.

A closer estimate of the full covariance and the resulting position-error ellipse, labeled "three-element diagonalized covariance in ($\overline{V}_{bo}$, in-track, cross-track) frame" (where $\overline{V}_{bo}$ is the velocity vector at burnout), is also shown in Figures 9 and 10. This covariance is generated by diagonalizing the covariance matrix at the estimated burnout time, setting the position variances equal to the satellite measurement variance (sigma squared), and regenerating the covariance using a reference frame computed using the estimated burnout-velocity vector. Propagating only the variances from burnout to reentry results in small variations from propagation of the complete covariance. This procedure yields a position ellipsoid that is accurate enough to generate an effective search pattern, because the difference between the full covariance and the three-element covariance is generally small compared to the size of one radar dwell. The reference frame based on the burnout velocity vector is a fair approximation for two reasons: most ballistic missiles are thrust in a direction close to the direction of their velocity vector, and the undetected acceleration between last track and burnout

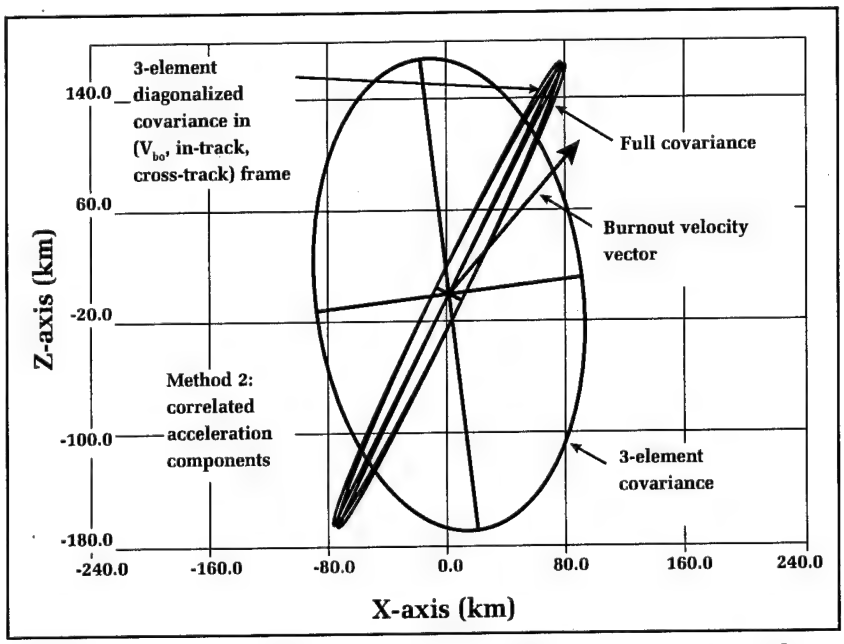


Figure 9. *Position ellipses at reentry, TBM-2 in X-Z plane, update interval 10 seconds, two satellites on X and Y axes, 13 simultaneous fixes, satellite estimated burnout at 135.0 seconds.*

tends to dominate other error sources.

Conclusions that can be drawn are summarized as follows:

- Satellites should transmit time, $\overline{R}$, $\overline{V}$, and three covariance parameters.
- Method 1: 3 σ's in Polar Geocentric frame
- Method 2: 3 σ's in $\overline{V}_{bo}$, cross-track, in-track frame
- Burnout time uncertainty (its effect on velocity sigmas) tends to dominate the other error sources.
- No need to transmit position sigmas for satellite covariance.

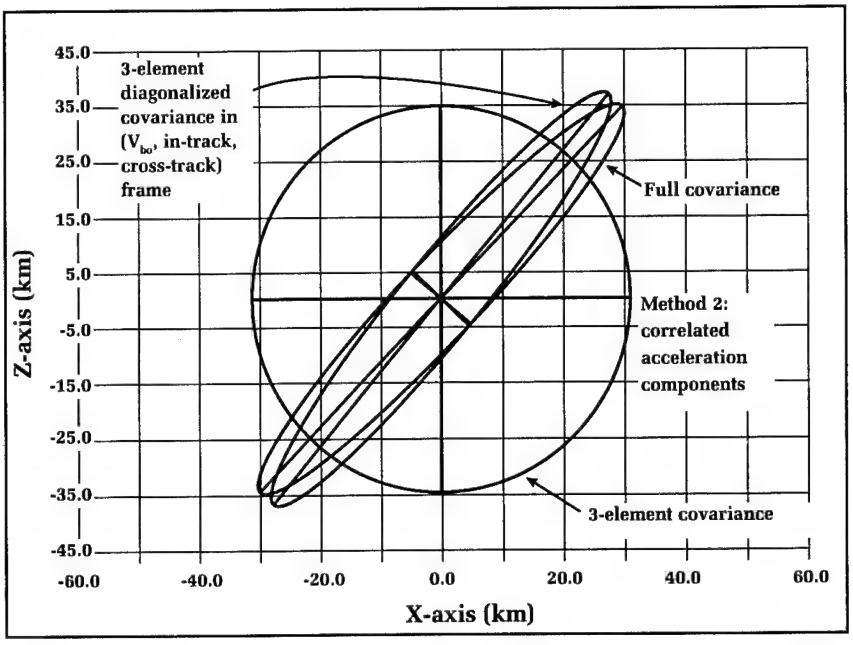


Figure 10. *Position ellipses at reentry, TBM-1 in X-Z plane, update interval 10 seconds, two satellitess on X and Y axes, 8 simultaneous fixes, burnout at 85 seconds.*

Ship/Surface Radar Tracking

Radar tracking of TBMs during launch and reentry is also plausible; the primary advantage is their ability to track during and after burnout. Actually, enough data to estimate the target's state vector can be obtained during a brief tracking period after burnout. The disadvantages of these active trackers are their power requirement and limited range to the target.

Tracking from surface ships is depicted in Figure 11, where active radar sensors are employed. Hand-off from the tracking ship near the launch area to the defending ship (or other ATBM platform) near the impact area is represented. The radar tracking measurement was simulated as the complete three-dimensional position error.

Only method 2 equations were used in the radar tracking study. Radar parameters are shown in Figure 12.

Radar tracking from surface ships shows significantly different results from satellite tracking (Figure 13). The radar-tracking device causes the position ellipsoid to be "wafer-shaped" and oriented such that the range error axis is roughly parallel to the last tracking LOS vector. The ellipsoid shown in Figure 13 appears to be almost flat because of the range accuracy of the tracking radar. This type of covariance matrix cannot be

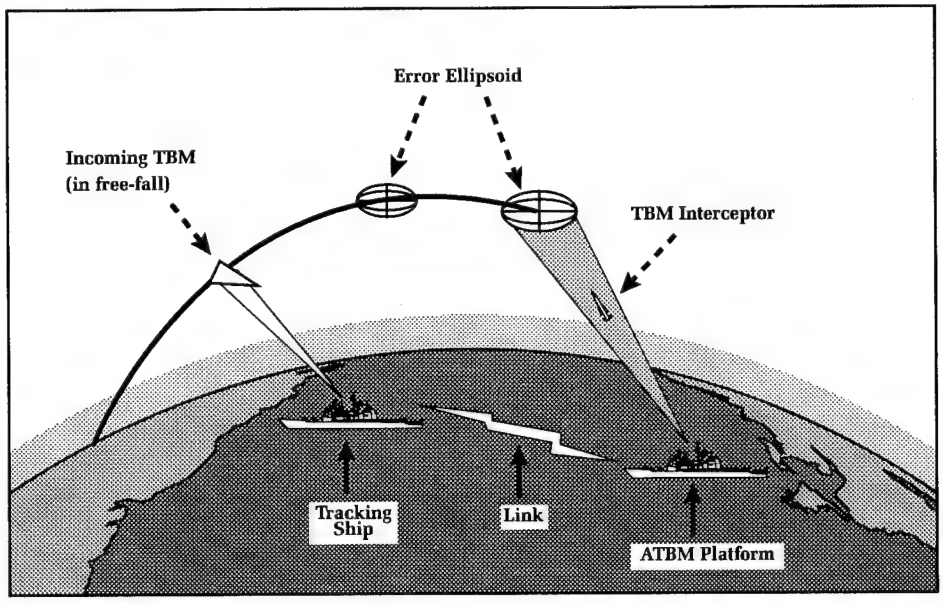


Figure 11. *Ship radar tracking and covariance hand-off study.*

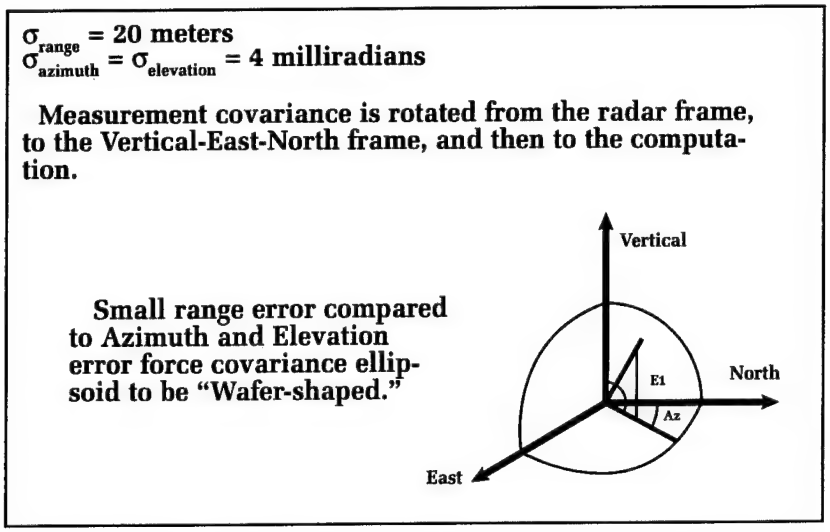


Figure 12. *Ship radar tracking parameters.*

accurately transmitted with the procedures described earlier for IR satellite tracking. However, velocity errors still dominate the shape of the ellipsoid; the problem here is the reference frame. In this case, a reasonable covariance can be obtained by transmitting a 3 by 3 velocity covariance and zeroing out the position errors (at last track).

Communication of Tracking Data

Communication of state vectors and covariance data from sensors to communication links to ATBM launch platforms is the conduit through which TBMD is achieved. The data has a very brief half life of value, and the conduit has a throughput limit. Hence, we are faced with the question of how can the input be reduced without destroying the output. Obviously, the state vector (time, position, and velocity) is the most important. If the target is maneuvering, then frequent updates are required. If the target is in free-fall, few updates to the track are needed.

Uncertainty in estimated position and velocity of the TBM is represented by the covariance matrix. The variances and correlation coefficients define the shape and orientation of the error ellipse. Correlation coefficients are zero when the error ellipse is symmetric about the coordinate frame it is written in. If the shape is skewed, then the variances and orientation data are needed to define the ellipse. The ultimate use of the covariance is to construct a radar search pattern (Figure 14). The ellipse shown here is the ellipsoid from Figure 13 projected into an acquisition radar frame centered 50 kilometers uprange from impact in the plane of the trajectory. Using a search pattern that covers the ellipse (size, shape, and orientation) will conserve valuable radar resources by limiting the search to only the most probable areas of the sky.

A sensitivity study was performed to determine the elements of the target-state covariance matrix that need to be transmitted to the shooter. IR satellites can track only during boost and will then relay trajectory state information at the estimated time of burnout. An ATBM platform near the impact area will receive this information and propagate it to the extent of its radar (start of acquisition). Propagation of the covariance from burnout to radar acquisition near reentry is accomplished via the transition matrix (shown in Figure 15).

These equations assume the TBM is in free-fall. Generally, TBMs free-fall for most

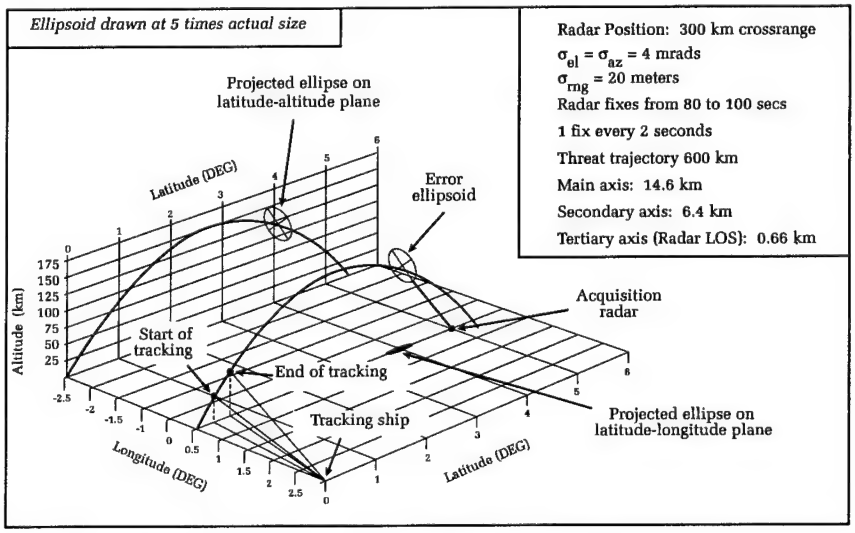


Figure 13. *Radar error ellipse (shown at reentry).*

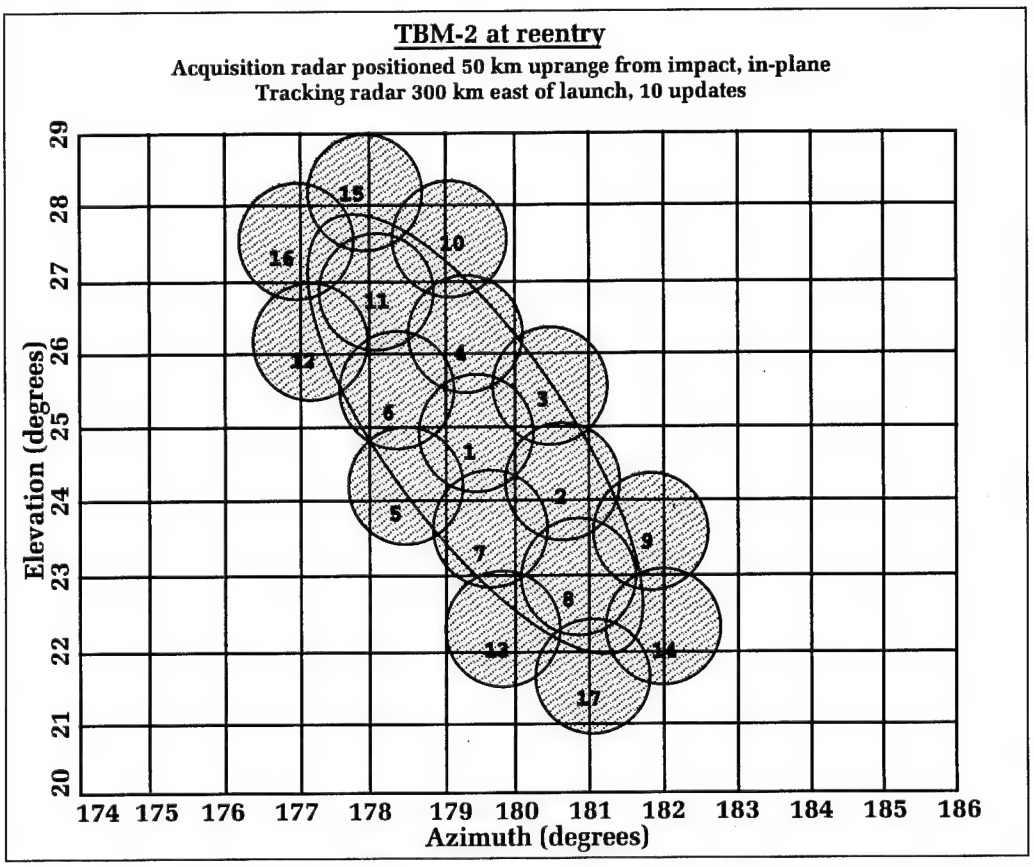


Figure 14. *Radar search pattern where shaded circles represent radar dwells. (Numbers indicate dwell placement order.)*

of their trajectory (that is why they are called ballistic). For satellite tracking, velocity errors grow over this time period and dominate any initial position error. Therefore, if the six elements of the upper triangular portion of the 3 by 3 velocity covariance (covariance matrices are symmetric about their diagonal) are transmitted and propagated to acquisition, the resulting position-error ellipsoid will be fairly close to the full covariance.

In the case of radar-to-radar cuing, if the state and state-covariance information must be propagated over a long period of time, again, velocity errors dominate. However, radar can track after burnout, possibly resulting in shorter times between cue and start of acquisition. In the cases where the velocity errors have not propagated long enough to dominate the position errors, the total error ellipsoid is very small and will fit in one radar dwell. Therefore, when cue-to-acquisition times are short, the six-element velocity covariance will still suffice, because the covariance really is not needed in these cases.

Trajectory State Propagation

The acquisition radar will receive target data via a communication link. The data will represent target-state estimates and state uncertainty at burnout or time of last track. The ship will propagate these data to a point

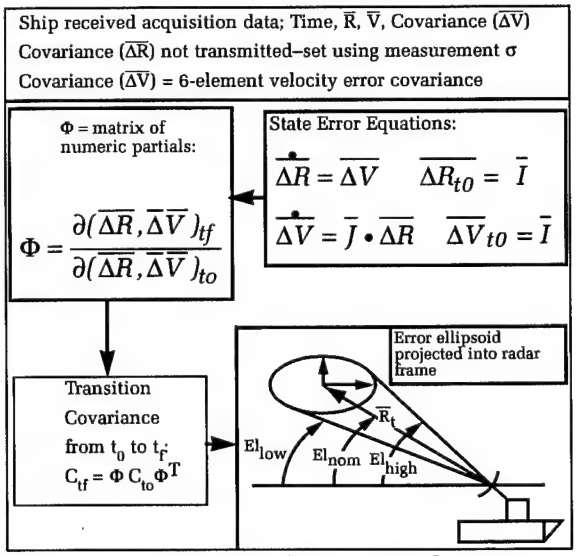


Figure 15. *Ship use of transmitted covariance data.*

where it can begin to acquire the target (Figure 16). The trajectory model used to propagate the state should be of a fidelity consistent with the uncertainty in the initial state.

A sensitivity study was performed for short- and long-range trajectories (Figure 17) to determine model complexity needed to propagate state information from the time of cue to the time of acquisition. A truth-trajectory model was defined to include the highest order gravity field, a fourth-order integration routine, standard atmosphere, and truth-aerodynamic data.

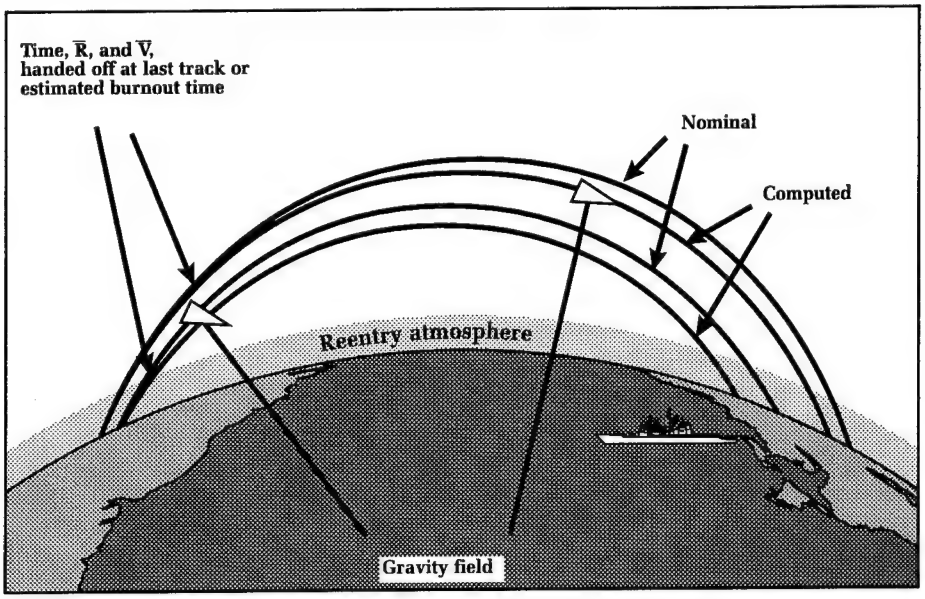


Figure 16. *TBM trajectory trade-off study.*

Bulleted items in Figure 18 make up the nominal truth-trajectory model. In the study, each parameter was tested independently while holding all other settings at their nominal value. The nominal trajectory was run to a stopping altitude of 10 kilometers, and the corresponding time was used to stop the perturbed trajectories, creating a three-dimensional position error. Two things were traded off in this study: computational speed and model accuracy relative to the truth model.

Errors resulting in reducing the complexity of the nominal model were computed and are depicted in Figures 19 and 20. Test cases include 9 launch points and 8 flyout azimuths, generating a total of 72 trajectories for each TBM type. In the two figures, the height of each bar represents the average RSS position error over all test cases. The width of each bar, excluding the ballistic coefficient sensitivities, represents the increase in computation speed over the nominal. The area of each bar can then be used to gauge each model's performance. The ballistic coefficient bars show the RSS position error incurred by errors in the ballistic coefficient of 10, 20, 30, and 40 percent. The width of the ballistic coefficient bars are not shown to scale, because they do not represent a reduction in model complexity and, therefore, do not increase computation speed. The nominal bar is shown as a baseline. The computer platform used to time these trajectory runs was a Sun Sparcstation 1+.

Figure 19 shows that, for TBM-1, using a flat-earth gravity field instead of the nominal field will increase the execution speed by 6.5 times but will also increase position error by 1.1 kilometers.

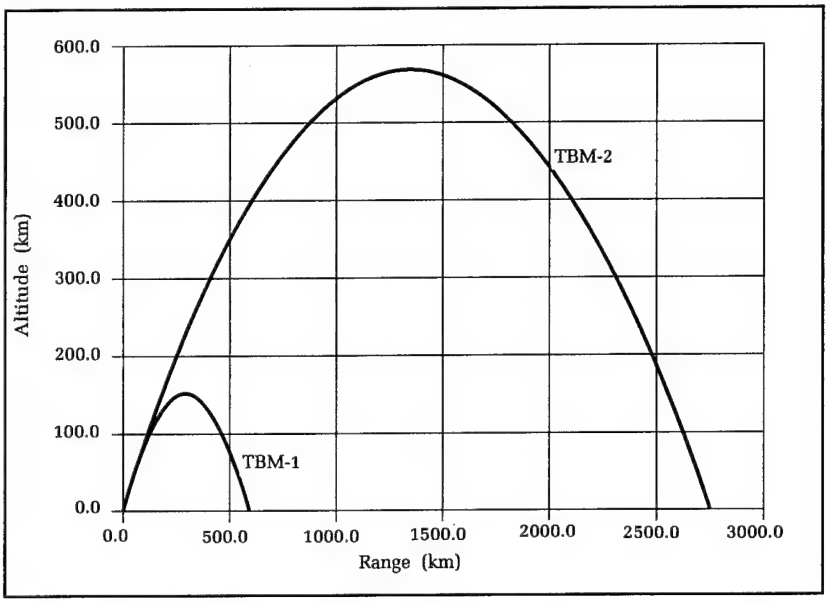


Figure 17. *TBM (TBM-1 and TBM-2) trajectories.*

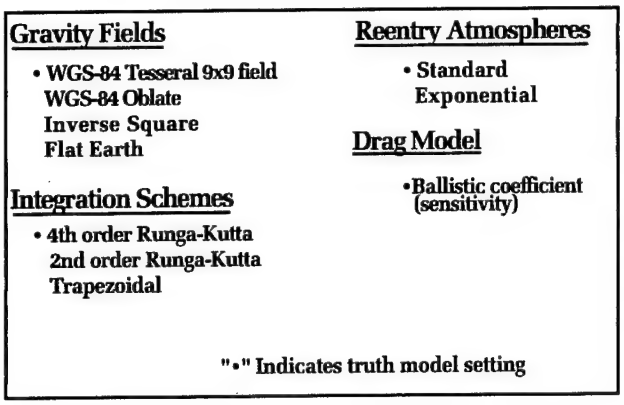


Figure 18. *Trajectory trade-off parameters.*

Using a round-earth gravity model will cause less position error (0.84 kilometer) but will run a little slower (6.4 times faster than nominal). The oblate gravity model will cause only 10 meters of error and will increase execution speed by 6.1 times. As expected, the second-order Runga-Kutta integration scheme ran twice as fast as the nominal fourth-order Runga-Kutta and caused 20 meters of error. The trapezoided integration scheme also ran twice as fast but introduced 50 meters of error. The exponential atmosphere approximation ran slightly faster than nominal but introduced over 2 kilometers of error.

The ballistic coefficient sensitivities show that if the estimated value of the ballistic coefficient is in error by 40 percent, 0.84 kilometer of position error will be incurred. This value is very large compared with the oblate gravity-field error and the integration-scheme errors. It is reasonable to assume that the ballistic coefficient may not be known very accurately. **Therefore, using an oblate gravity field with any of the integration schemes studied in the trajectory propagation model would seem reasonable.**

The exponential atmosphere approximation produced a marginal increase in execution speed and introduced considerable error. **Therefore, using an exponential atmosphere in the trajectory propagation model should not be considered.**

Figure 20 shows the trajectory model performance for the longer range TBM-2. The flat-earth and round-earth gravity models produce much larger errors for the longer trajectory. However, the conclusions that can be drawn are the same as for the short-range case.

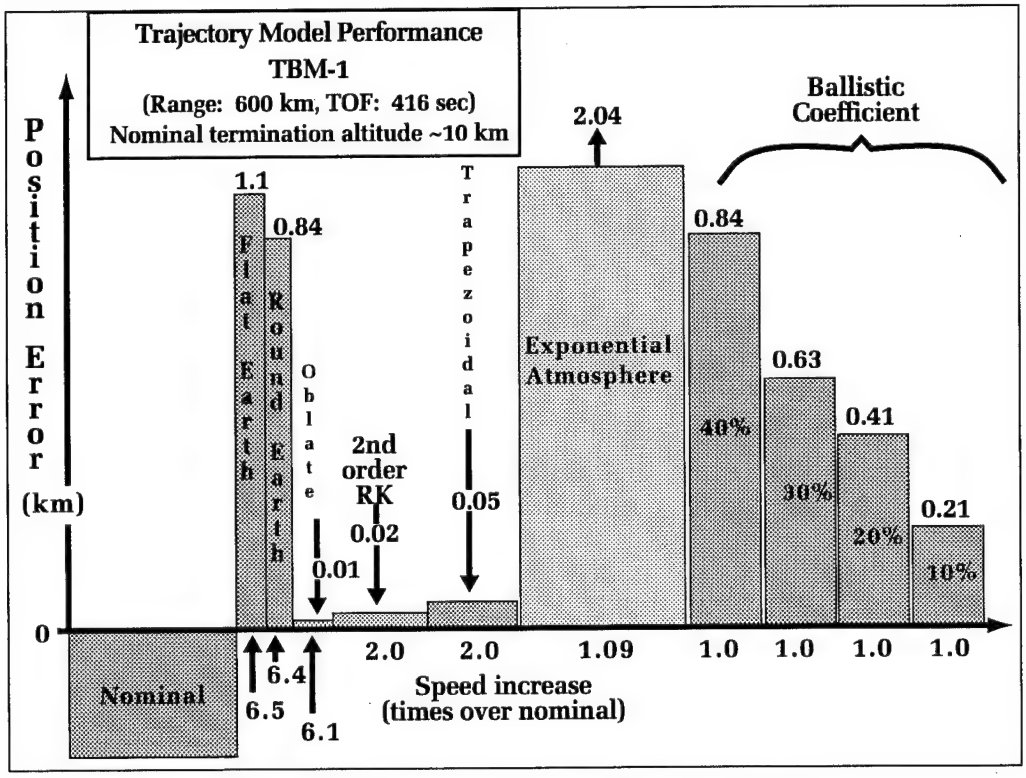


Figure 19. *Trajectory model performance TBM-1.*

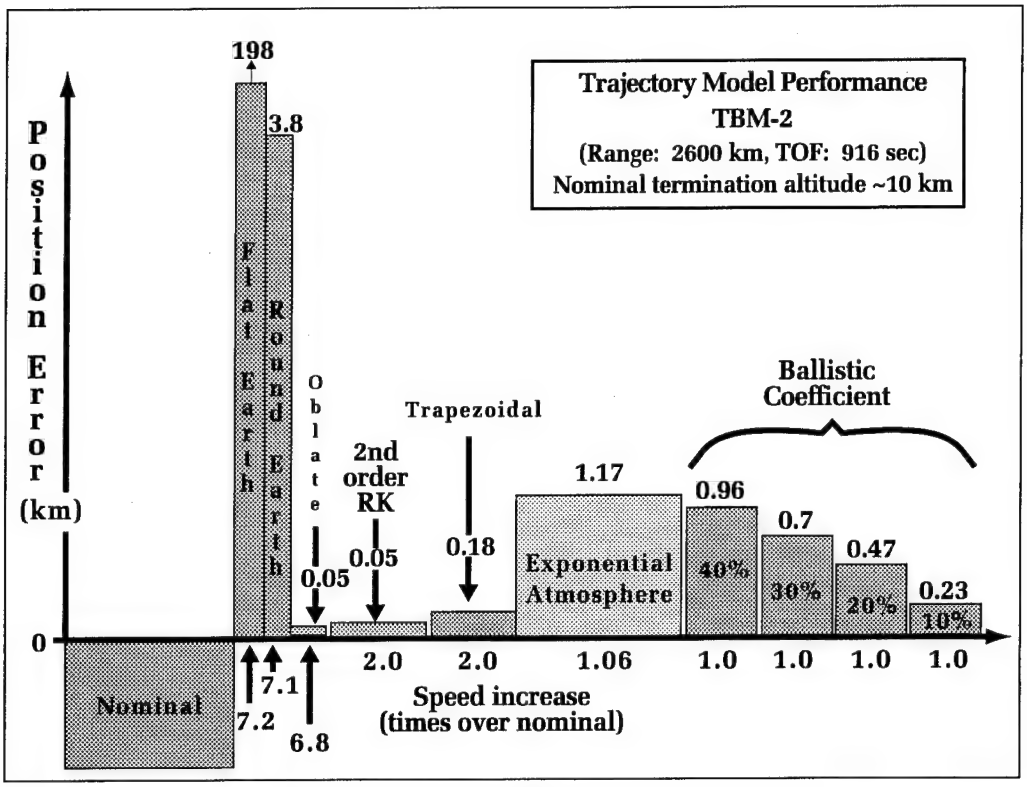


Figure 20. *Trajectory model performance TBM-2.*

The data in Figures 19 and 20 were derived using a constant integration step size of one second. Integration step size was also varied in the study to determine how the various integration schemes would perform. The fastest execution time that achieved reasonable accuracy was obtained using the fourth-order Runga-Kutta integration scheme with two different step sizes. When the TBM was coasting through space, a large time step was used. When the TBM started to reenter the atmosphere, a smaller time step was used. This process resulted in execution speeds up to nine times faster than the constant 1-second time step, while introducing less than 10 meters of error.

Conclusions

Space/ship tracking of TBMs is practical and efficient. Cuing data consists of target state and a "subset" of the covariance matrix. Covariance data is ultimately needed to generate a search pattern for the shooter radar and, as a result, can be considerably reduced in size with little or no effect on system performance. Trajectory model fidelity for propagation of the target-state vector will be selected based on availability of computer resources and initial-state vector fidelity.

Acknowledgment

The authors would like to acknowledge John Lawton for his input and assistance on this project.

Reference

1. Danis, N. J., "Spaced-Based Tactical Ballistic Missile Launch Parameter Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 29, No. 2, Apr 1993, pp. 412-424.

The Authors

GREGORY DRESCHER has worked as a physicist in the SLBM Systems Accuracy Branch at NSWCDD since June 1985. His main areas of expertise include post-flight error analysis and trajectory reconstruction of Trident II missile test flights. He graduated from Slippery Rock University of Pennsylvania in 1985 with a B.S. in physics. In 1988 he received a patent, and in 1993 he received a Strategic and Space Systems Department technical excellence award and Tactical Ballistic Missile group achievement award.

GARY SITZMAN is a senior analyst at NSWCDD. He has written numerous technical reports in the areas of missile trajectory simulation, system accuracy, and gravity modeling algorithms. Currently, his area of work involves developing new applications of existing strategic missiles. He graduated from Westmar College in 1964 with a B.S. in mathematics. He later received an M.S. from the University of South Dakota and came to NSWCDD in November of 1966. In December of 1992 he was the recipient of the Navy Meritorious Civil Service Award.

Superconducting Magnetic Sensors for Mine Countermeasures

Ted R. Clem

Operation Desert Storm clearly showed the importance of mine countermeasures (MCM) in protecting U.S. naval assets. Sensor development for mine detection represents an important area for MCM. The Naval Surface Warfare Center (NSWC) Coastal Systems Station (CSS) developed a ruggedized, superconducting magnetic sensor capable of operating on board undersea and airborne towed platforms. With this sensor, CSS was able to demonstrate buried mine detection for the U.S. Navy. Subsequently, the sensor was incorporated into a multisensor underwater towed vehicle to provide a robust mine-hunting capability. This article describes the superconducting sensor, its role in mine hunting, and recent technology developments, including progress with the new high-critical-temperature (T_c) superconductors. The technology also has dual-use applications, including environmental cleanup and biomedical examination.

Introduction

A superconductor is a material that displays unusual behavior at low temperatures, notably the absence of electrical resistance.[1,2] In addition, a superconductor has unusual magnetic characteristics, in particular, perfect diamagnetic behavior; i.e., the total exclusion of magnetic flux from the interior of the superconducting material. There is a T_c below which the material exhibits these superconducting characteristics and above which the material typically exhibits classical metallic, insulator, or semiconductor behavior (but not the unusual superconducting properties). Prior to 1986, the known superconductors exhibited critical temperatures less than 25 degrees Kelvin (K), notably a T_c of 9 K for niobium. Liquid helium cooling to 4 K has been required to induce the superconducting state in this class of superconductors. Starting in 1986, new classes of superconducting materials were discovered with critical temperatures exceeding 77 K, the boiling point of liquid nitrogen.[3,4] This offered new opportunities for naval applications, largely as a result of reduced refrigeration requirements, which permit more practical sensors for Fleet operations.

Superconducting sensors provide the greatest sensitivity available today for magnetic anomaly detection. During the 1980s, CSS developed a superconducting magnetic sensor specifically dedicated for deployment in an undersea towed vehicle, demonstrating for the first time high sensitivity coupled with rugged, robust, and reliable performance outside the laboratory environment. The sensor employed niobium superconducting technology cooled with liquid helium. It was used to demonstrate buried mine detection for the U.S. Navy.

Subsequently, a mine-hunting system concept was developed during the mid 1980s under the Magnetic and Acoustic Detection of Mines (MADOM) Project to more robustly address the unusual difficulties and complexities associated with the buried mine problem. The MADOM concept employed the fusion of a triad of sensors for mine detection (the superconducting magnetic sensor described above, along with two sonars) operating from an underwater towed-vehicle system. The triad of sensors was also used to help distinguish between mines and mine-like objects. This article will describe the superconducting magnetic sensor along with its role in the MADOM mine-hunting concept. In addition, it will address advances in this sensor technology, especially opportunities associated with high T_c technology using liquid nitrogen cooling in place of liquid helium cooling.

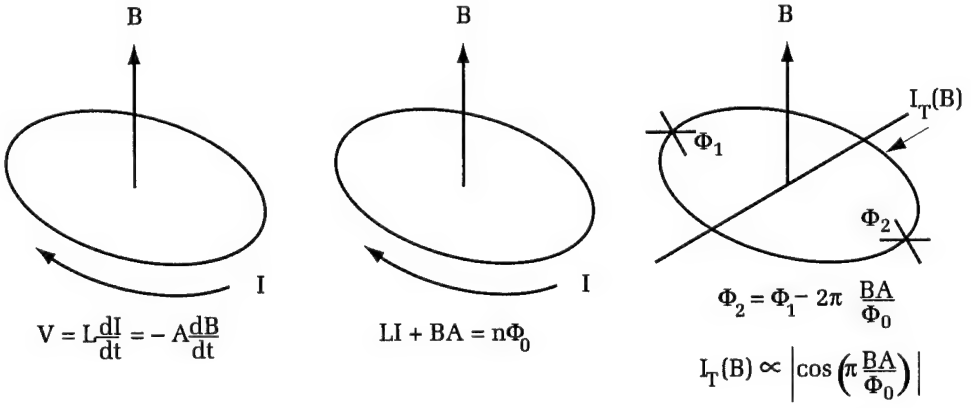
Superconducting Magnetic Sensors

In order to establish a rudimentary appreciation of the physics underlying this sensor technology, consider a loop incorporating a normal metal conductor, as displayed in Figure 1(a). Faraday's Law states that transient voltages and currents are established in this loop as a result of a change in flux. This effect explains the alternating current flux-to-voltage transformation that underlies the operation of a conventional induction-coil magnetic sensor. Since the voltage increases with frequency, this type of sensor is most effective at higher frequencies.

For a closed superconducting loop, as displayed in Figure 1(b), there is a similar effect that permits direct current (dc) and low-frequency, flux-to-current transformations.[1,2] Coherent macroscopic quantum states can be established around the loop, an effect that has been readily demonstrated for wire and thin-film loops, with dimensions on the order of one meter. As a result of this macroscopic quantum effect, magnetic flux is quantized inside the loop. Moreover, persistent currents are established as a result of a flux change, in contrast to the transient voltages and currents established for normal conductors. Hence, we have a dc flux-to-current transformer providing the basis for a low-frequency, magnetometer-sensing loop.

Brian Josephson predicted the existence of quantum tunneling of superconducting electron pairs across a barrier, maintaining coherence of a wave function across the barrier. He also predicted that voltages will be established across a barrier as a result of changes in phase across the junction.[5] This phenomenon was subsequently demonstrated by P. W. Anderson and J. M. Rowell.[6] It provides the basis for operating a Superconducting Quantum Interference Device (SQUID), the key element for the class of magnetic sensors under discussion.[1,2,7]

The dc SQUID is a closed superconducting loop, as depicted in Figure 1(c), broken by two barriers that are referred to as *Josephson junctions* (as depicted by Xs in the figure). As a consequence of the existence of a macroscopic wave function around the SQUID loop, the phase across the junctions will change in response to flux changes inside the loop. This, in turn, will induce corresponding voltage changes in the junctions, as predicted by Josephson, and will



(a) Normal Metal Loop (b) Superconducting Loop (c) dc SQUID

Figure 1. *Equations relating the voltages, V; current, I; and quantum phase, Φ, induced by changes in magnetic induction, B, for several configurations of interest. (In this figure, A denotes the area of the loop, and L denotes its inductance. Φ_0 denotes a unit of magnetic flux with magnitude of 2.07 x 10^{-15} Webers.)*

modulate current flow through the SQUID. Such devices provide the most sensitive magnetic field detection that has ever been demonstrated in the low-frequency regime (<100 Hz).

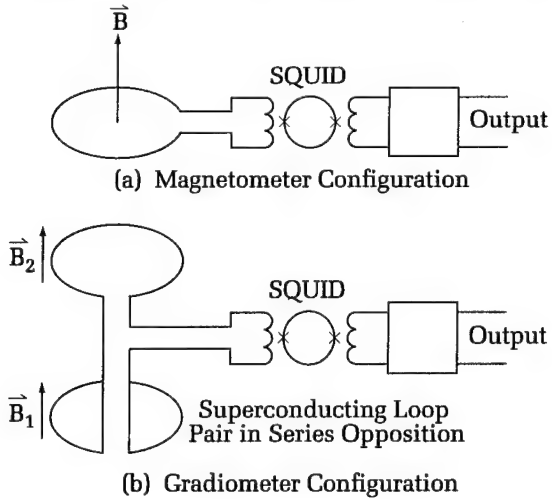
Typically, a SQUID is connected to a low-frequency superconducting antenna, such as the magnetometer sensing loop displayed in Figure 1(b), in order to increase the gain of the circuit. Such a magnetometer is shown in Figure 2(a). For other applications, a gradiometer loop is used to measure spatial differences of field in order to compensate for the very large effects of the Earth's field (Figure 2(b)).

Superconducting Gradiometers Outside the Laboratory Environment

Essentially all the efforts with SQUID sensor technology dealt with sensors inside a very controlled laboratory environment. Laboratory SQUIDs provide uncontested sensitivity over rival technologies and, as demonstrated, they provide superior capability in field deployment even on board moving undersea and airborne platforms.

As first shown both theoretically and in the field, superconducting gradiometers can efficiently and accurately determine the range and bearing of a magnetic target, such as a mine or a submarine, and identify the target in terms of its magnetic characteristics. For this type of application, especially for operation on board a moving platform, a gradiometer to measure spatial differences of field is very desirable for common mode rejection of the very large signals generated by the Earth's field. Small dimensions imply extreme coherence in motion because of the small baselines involved, as compared to more bulky or less sensitive technologies where longer baselines are used. In addition, with the compact sizes, we can pack a large number of gradient channels into a small package and gain multiple observations of a target. Figure 3(a) shows one relatively simple means to package five gradient channels (sufficient to characterize the gradient tensor field), along with three magnetometers for motion compensation. With such a configuration, the gradiometer collects the time history of the five gradient and three field components of the magnetic field in its frame of reference. A typical time history for a gradiometer moving past a magnetic dipole source is displayed in Figure 3(b). This data-collection approach permits complete characterization of magnetic dipole sources at long ranges, obtaining localization and classification information not possible by other magnetic sensor technologies incorporating only a single magnetometer channel, such as the U.S. Navy's AN/ASQ-81, an optically pumped total-field magnetometer.

CSS has actively investigated sources of noise that could potentially deteriorate gradiometer performance. Some noise sources occur when the sensor is operated "unshielded," even when the sensor is stationary; e.g., temporal changes in the Earth's magnetic field and electromagnetic interference. Other noise sources are associated with the sensor's motion (e.g., relative motion of the sensor with respect to the geological structure, ocean waves, and the sensor's platform) and intrinsic changes internal to the sensor. Compensation approaches have been investigated and implemented in order to minimize the effects of such noise sources.



(a) Magnetometer Configuration

(b) Gradiometer Configuration

Figure 2. *Conceptual sketches of superconducting SQUID-based magnetic sensors.*

Principles of Gradiometer Detection, Classification, and Localization

The performance of a gradiometer is measured by its detection range, which is a function of the sensitivity of the gradiometer channels in motion and the magnetic moment of the targets of interest. A magnetic dipole is fully characterized by its magnetic dipole moment, $\vec{m}$ (in meter-kilogram-second units here of tesla meters cubed ($T \cdot m^3$)). The magnetic field induction, $\vec{B}$, generated by the magnetic dipole is the vector, and the spatial gradient of the magnetic field induction is a second-order tensor with components given by

Glossary

- The unusual properties of superconductors arise from *electron pairing*. In the superconducting state, an attractive potential is established between pairs of electrons, and it becomes meaningful to define quantum wave functions for these pairs. For normal metals, resistance is attributed to single-electron scattering arising from electron-phonon interaction. Such electron scattering is eliminated for the electron pairs in a superconductor, and we have the explanation of resistless current flow in a superconductor. For the conventional superconductors discovered prior to 1986, with few exceptions, the electron pairing can be explained by electron-phonon interaction. At the time of this writing in 1994, intense theoretical investigations continue to explain the mechanism underlying electron pairing for the new high-temperature superconductors.

- In quantum mechanics, *tunneling* is a phenomenon in which quantum particles can pass through a potential barrier of microscopic width, although the corresponding classical particle would not pass through the barrier. The quantum particle is characterized by a probability wave function that has spatial extent so that there is a certain probability for the particle to pass through the barrier.

- A *magnetometer* is a sensor that measures the absolute magnetic field or changes in magnetic field as a function of position or time. A magnetometer may measure vector components of the magnetic field or the magnitude of the field.

- An *optically pumped magnetometer* measures the wavelength of an element's spectral lines in a magnetic field in the Zeeman effect. In this sensor concept, the vapor of an element such as Cs or He-4 is optically pumped with circularly polarized light causing outer electron excitations. In addition, electron spin flipping is induced by a radio-frequency electromagnetic field. This combination of optical and radio-frequency effects creates a sharp resonance whose frequency is dependent on the magnitude of the magnetic field.

- A "magnetic" *gradiometer* measures a derivative of a magnetic field, either a vector component or the absolute magnitude, with respect to position. In this article, the discussion is limited to first-order derivatives, such as $\vec{m}\,\partial B_x/\partial y$. In practice, the magnetic-field component is measured at two distinct spatial points, and the derivative is estimated as the ratio of the difference of field over the spatial separation.

- A *magnetic dipole* is an electrically conductive loop carrying an electrical current. For the case in which the loop is circular, the *magnetic moment*, $\vec{M}$, of the dipole is defined as the product, $\vec{M} = IS\vec{n}$, where the magnitude of $\vec{M}$ is the product of current, I, and loop area, S, and $\vec{n}$ is the unit vector normal to loop surface. The magnetic dipole moment, $\vec{m}$, for the loop is given by the expression,

$$\vec{m} = \frac{\mu_0}{4\pi}\vec{M}$$

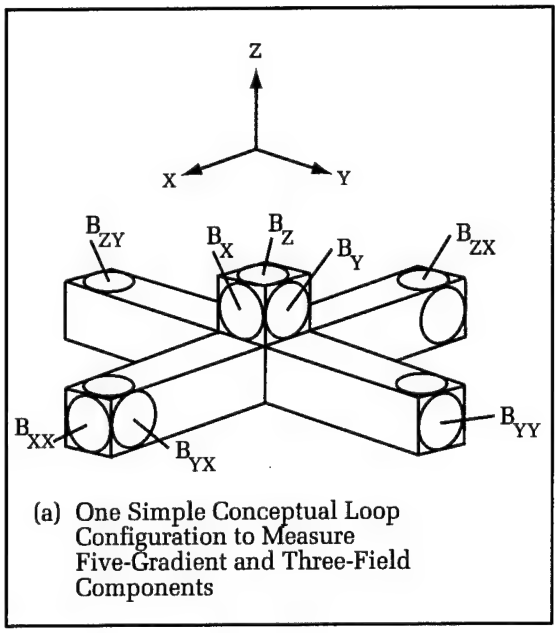
The magnetic field induction, $\vec{B}$, of the dipole is given in the far field by the expression

$$\vec{B} = \frac{3(\vec{m}\bullet\vec{R})\vec{R}}{R^5} - \frac{\vec{m}}{R^3}$$

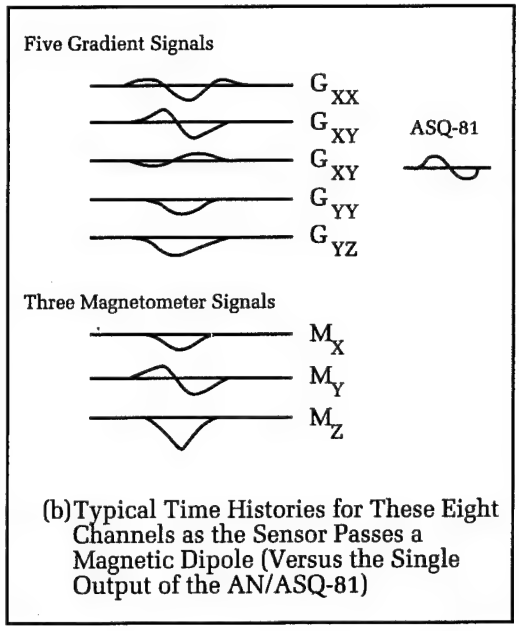
The concept of magnetic moment and *magnetic dipole moment* can be extended to describe general current distributions and magnetized bodies. In these cases, the magnetic field induction in the far field, away from the current distribution of the magnetized body, can be described by the same expression for $\vec{B}$ in terms of its generalized magnetic dipole moment.

- *Thin-film processes* developed for the silicon microelectronics industry have been applied to the superconducting technology. Common industrial design rules are established, in which superconducting niobium films are deposited and patterned reproducibly and accurately, with widths on the order of 2 microns (micron = 10^{-6} meters) and thicknesses on the order of 0.1 micron.

- At atmospheric pressure, helium liquifies at 4.2 K. A new "superfluid" thermodynamic phase occurs when helium is cooled down to the *lambda point*, approximately 2.2 K at atmospheric pressure. The new phase is characterized by the "super conduction" of heat. This property provides the basis for refrigeration approaches that are extraordinarily isothermal.



(a) One Simple Conceptual Loop Configuration to Measure Five-Gradient and Three-Field Components



(b) Typical Time Histories for These Eight Channels as the Sensor Passes a Magnetic Dipole (Versus the Single Output of the AN/ASQ-81)

Figure 3. *Compact multichannel sensor packages can be implemented with this technology to provide detailed localization and classification capabilities.*

$$\vec{B} = \frac{3(\vec{m} \bullet \vec{R})\vec{R}}{R^5} - \frac{\vec{m}}{R^3}$$

The spatial gradient of the magnetic field induction is a second-order tensor with components given by

$$B_{ij} = -\frac{3}{R^7} \bullet \left\{ \vec{m} \bullet \vec{R}(5R_iR_j - R^2\delta_{ij}) - R^2(R_im_j + R_jm_i) \right\}$$

As a result of Maxwell's equations, only five of these nine tensor elements are independent. For this reason, gradiometers are typically designed with five independent gradient channels using the minimum number that permits characterization of the local tensor gradient field.

In the naval applications of mine and submarine hunting, the gradiometer is operated on board a platform moving past a magnetic target, such as a mine. By means of sampling data over a period of time, the locations and the magnetic moments of multiple targets can be determined. Typical gradient signals observed by a five-gradient channel sensor passing by a magnetic target are displayed in Figure 3(b).

The Field Deployable Superconducting Gradiometer Package

This section will describe the basic subassemblies and critical components of the Superconducting Gradiometer/Magnetometer Sensor (SGMS), the magnetic sensor used for demonstrating buried mine detection. This sensor incorporates conventional low T_c materials, such as niobium, and is cooled by a liquid helium. An artist's conception of this system is shown in Figure 4. It has three major subassemblies: (1) a passive cooling unit referred to as a dewar, (2) a probe assembly, and (3) a room-temperature sensor electronics package.

In order to provide optimal performance in motion, the SGMS dewar (shown in Figures 4 and 5(a)) has a horizontal design, distinguishing it from standard laboratory dewar designs. Special designs and stringent materials selection (not necessary for laboratory applications) are required to assure sensor performance in motion. The dewar is 44 inches long and 13 inches in diameter. The capacity of the dewar is 15 liters of liquid helium, with a hold time of three days. The dewar has four concentric chambers. The innermost chamber houses the probe assembly, the center chamber holds liquid helium, and the outermost chambers are evacuated, with the walls between them serving as a thermal shield. The probe assembly is cooled by evaporated helium gas, which is vented to the atmosphere.

The probe structure, shown in Figures 4 and 5(b), fits into the innermost chamber of the dewar and is composed of two structural components—a neck plug and a gradiometer support structure. The neck plug provides the mount for the probe structure into the dewar and provides thermal insulation for the open end of the dewar. The gradiometer

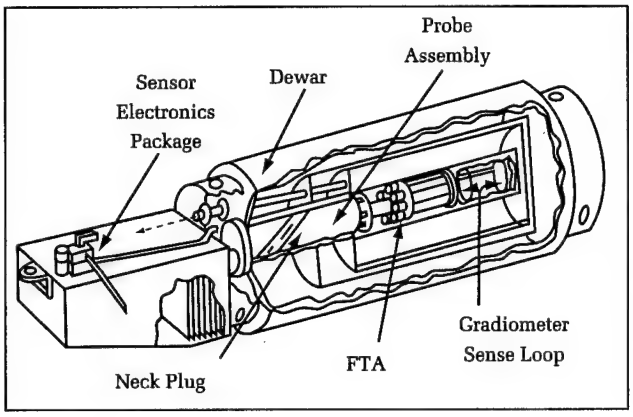


Figure 4. *An artist's conception of the SGMS.*

support structure is a combination of structural cylinders and flanges on which the SQUID assemblies, known as Flux Transformer Assemblies (FTAs), and the sense-loop array are mounted. For magnetic target detection and localization, the SGMS incorporates five gradient sense loops (to determine all components of the local tensor gradient field) and three additional magnetic field sense loops (for motion compensation of the gradiometer signals).

The FTA, conceptually depicted in Figure 6, is the heart of the sensor. It houses the SQUID and the transformers that interface the SQUID to its sense loop and to its warm-side electronics. The FTA components are mounted in a superconducting canister for shielding against electromagnetic interference and external magnetic fields. Modulated analog SQUID output signals are coupled electrically via transmission lines to the room-temperature electronics package outside the probe/refrigeration assembly.

The electronics package for the SGMS permits signal readout and circuit control for each of the eight channels. The package consists of SQUID electronics, antialiasing filters, analog-to-digital converters, and control circuits for remote tune, gain, and offset control. Digitized signals are transmitted electrically to a computer for data processing.

Two key points should be made about this generation of sensor. First, the sensor employed bulk niobium superconducting components for the SQUID and niobium wire-wound superconducting loops. Second, it used a liquid helium dewar to cool the components down to 4 to 5 K in order for the components to superconduct.

The Magnetic and Acoustic Detection of Mines

Generally, active acoustic approaches have proven to be an effective means to detect, classify, and localize tethered sea mines or bottom mines exposed (proud) with respect to the bottom. However, the shallow-water buried mine environment is an especially difficult acoustic environment in which to operate. Interfering reverberations from the air/sea and sea/bottom interfaces, bottom topographical features, general harbor debris, and the problem of penetrating the sea floor present a difficult acoustic environment for buried mine detection.

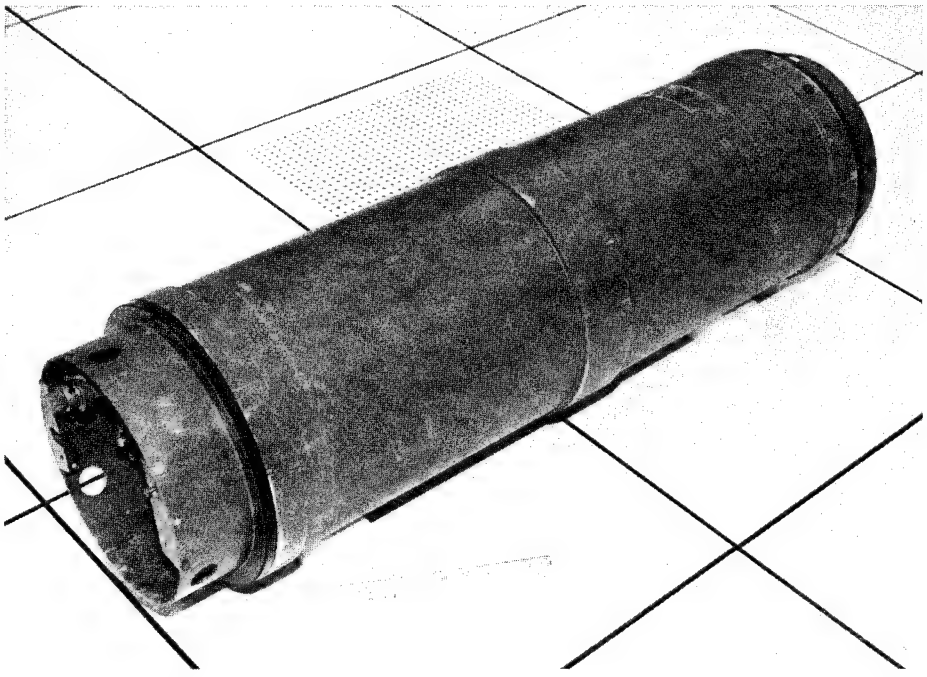


Figure 5(a). Dewar.

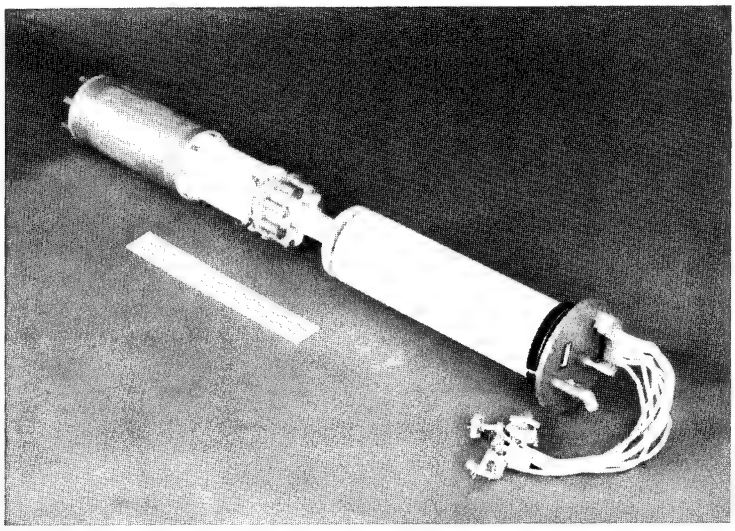


Figure 5(b). *Major subassemblies in the SGMS package.*

In contrast, magnetic sensors are not limited by the nature of the ocean bottom or surface. Consequently, they are capable of detecting ferrous bottom mines proud, partially buried and, most significantly, totally buried. In addition, this sensor detection capability is omnidirectional, so that it is capable of localizing targets to the side, the front, above, or below the vehicle. Hence, it does not suffer from some of the limitations familiar to side-scan sonars, notably range compression and voids (holidays). This permits a detection capability against all mine types from fully buried bottom mines to tethered mines to floaters. As will be discussed subsequently in considerable detail, the fusion of magnetic and acoustic sensors provides a most significant means for clutter rejection.

Unisys[8] developed the SGMS system and provided the first two sensors in 1982. In 1983, after a series of successful technical tests, noise sources were identified that plagued the gradiometer when it was placed in motion typical of deployment. As a result of dedicated experimentation, the system was modified in 1984, leading to a dramatic increase in performance. More recently, in 1989 one of the SGMS systems was refurbished by IBM[9] and Quantum Design,[10] leading to impressive improvement in sensor reliability and operability.

Starting in 1984, the SGMS system was first operated on board an undersea vehicle towed from a surface ship. With this technology, CSS was able to perform buried mine detection. The SGMS system demonstrated, for the first time, high sensitivity and rugged, robust, reliable, and fully automated performance of such a sensor operating on board a moving, towed underwater vehicle outside the laboratory environment. In fact, the sensor was used for a period of 7 years for Exploratory Development (6.2) and Advanced Technology Development (6.3A) sea testing.

Following this initial success, CSS initiated the 6.2 MADOM Project starting in 1985. The SGMS was incorporated into a sensor suite, including a synthetic aperture sonar (SAS) and a high-frequency, side-scan sonar operating from towed underwater vehicles. This system concept is depicted in Figure 7. The side-scan sonar is housed in the forward underwater towed vehicle, while the SGMS system and the SAS are housed in the aft vehicle. A picture of the tow vehicle for the SGMS system and the SAS is shown in Figure 8. Running concurrently with the last

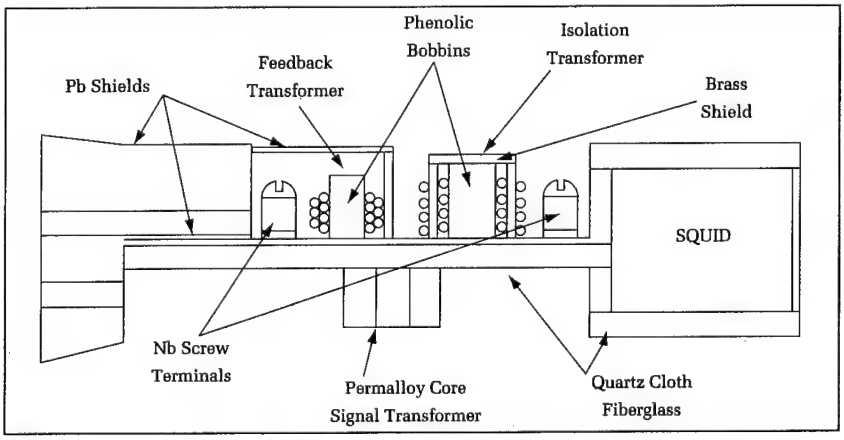


Figure 6. *Diagram of the FTA assembly.*

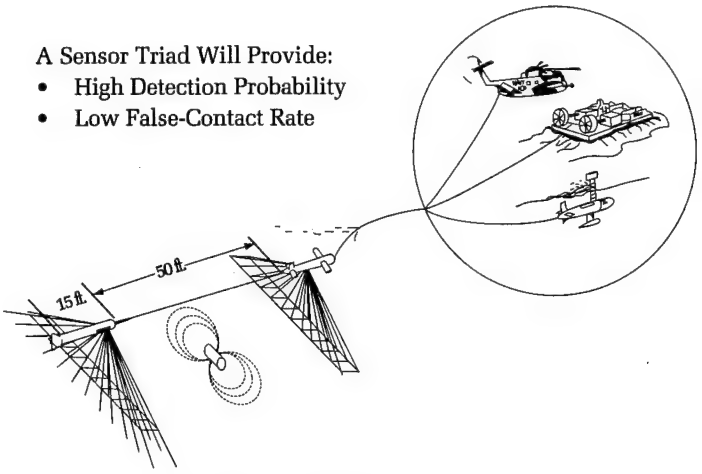


Figure 7. *The MADOM system concept.*

two years of the 6.2 Program, a 6.3A Advanced Technology Demonstration was successfully completed in 1990. This success proved a capability to detect and classify buried mines. This, in turn, led to the transition of the MADOM concept to a 6.3 Advanced Development Model phase being developed under the Buried-Mine Detector Project.

Parallel Signal and Image-Processing System

The Parallel Signal and Image-Processing System (PARSIPS) was developed for the MADOM Project in order to control the sensor suite and process data collected. The heart of the PARSIPS computer system is a bank of eight array processors—four of which are dedicated to sonar data conditioning, motion compensation, and beamforming; and four of which are dedicated to gradiometer operation (data conditioning, motion compensation, and target localization). Of particular interest in this article, PARSIPS processing of gradiometer data provides magnetic target detection and classification (D/C), as well as localization fully automated effectively in real time; i.e., the magnetics Computer Automated Detection/Classification (CAD/CAC) problem has been fully solved. In contrast, a parallel capability for acoustic CAD/CAC is currently under development and has not yet been fully demonstrated.

Through PARSIPS, the sensor triad is highly automated and is easily managed by a single person. In particular, gradiometer operation consists of turning on the power to the gradiometer and the helium pump, tuning the gradiometer (a 3-minute, semiautomated process typically executed at the start of operation), inputting towing vessel speed, and executing keystrokes to start and stop the automated gradiometer data recording and processing.

The gradiometer collects the time history of the gradient of the magnetic field in its frame of reference as it moves along its trajectory through the magnetic field. From this collected data, the gradiometer analysis algorithm provides real-time estimates of the locations and moments of the magnetic dipoles. This algorithm operates on overlapping segments of the gradiometer data, compensating for the motion of the sensor and iteratively solving for the three-dimensional position and moment vectors of the

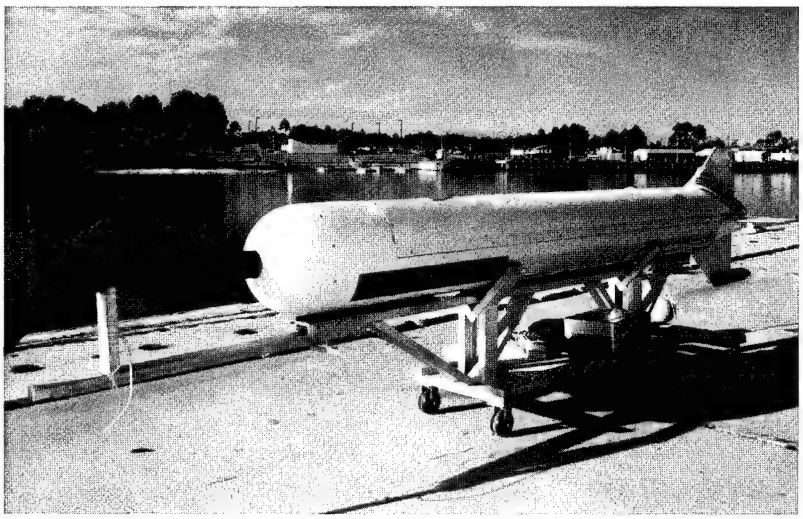


Figure 8. *The undersea tow vehicle for the SGMS and the SAS.*

source dipoles until a minimal set of dipoles is found that is sufficient to describe the measured signal. This magnitude of the magnetic moment of a target is used to classify that target as either a magnetically mine-like or non-mine-like object. The algorithm requires no inputs of any kind (such as thresholds or instrument noise levels), except for the tow speed.

The PARSIPS display is shown in Figure 9. The images of the sea floor obtained from the SAS and side-scan sonars are displayed waterfalling in real time on the left and right portions of the high-resolution monitor. The locations of magnetic targets found by the automated-gradiometer processing algorithm are displayed as icons superimposed on the sonar images of the bottom. The shape and color of these icons denote the size of the target's magnetic moment and the relative confidence of the algorithm in the target's localization. The icons are mapped onto their locations on the screen after a delay following the point of closest approach in order to assure sufficient data for the statistical analysis, depending on the density of magnetic targets.

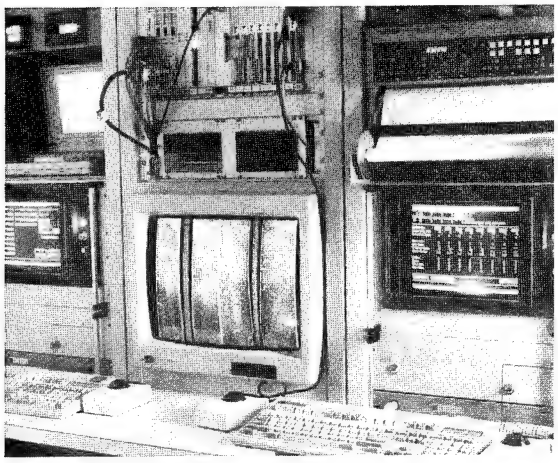


Figure 9. *The PARSIPS sonar/gradiometer display.*

Clutter Rejection

The combination of both acoustic and magnetic D/C has been demonstrated to successfully minimize the probability of false mine classification; i.e., the problem of clutter rejection. The sea bottom may be cluttered with debris that might be falsely identified as a mine. Figure 10 depicts the

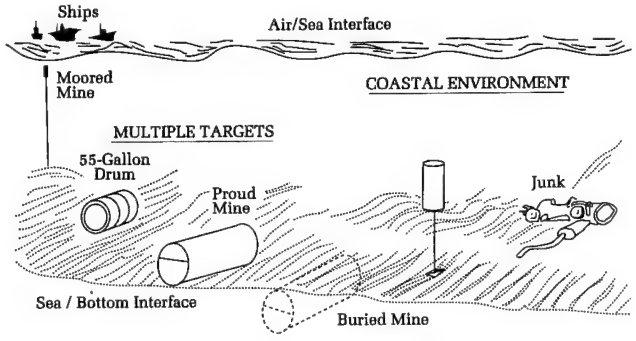


Figure 10. *A typical environment encountered in a bottom mine-hunting operation. The mine-hunting system must distinguish among mines, debris, and bottom features.*

environment in which a mine-hunting system may operate, displaying bottom mines and bottom debris, such as sunken cars and 55-gallon drums.

A magnetic sensor will identify a target as mine-like if the magnitude of the target's moment falls within a certain range of acceptable values. A target's magnetic moment is principally a function of the amount of steel in the mine. Hence, harbor debris such as 55-gallon drums and sunken automobiles will be classified as nonmine-like if their magnetic moments fall outside the acceptable range. However, other harbor debris such as car axles, engine blocks, and refrigerators will be magnetically indistinguishable from a ferrous mine if they have magnetic moments comparable to that of a ferrous mine.

Conversely, the acoustic sensors rely on the shape of an acoustic target in order to classify that target as mine-like or not. However, a mine's shape may be acoustically indistinguishable from the shapes of nonmine objects. Thus, the acoustic sensors would classify 55-gallon oil drums as mine-like, but not truck axles.

Fusion of the magnetic and acoustic data provides the means to distinguish both a 55-gallon oil drum and a truck axle as nonmine objects. From this line of reasoning, one can see that the fusion of magnetic and acoustic data provides a powerful tool for discriminating between mines and nonmines.

Thin-Film Sensor Development

The following stringent requirements must be satisfied by a field-deployable sensor operating on board a moving platform:

(1) unshielded operation in the Earth's magnetic field with rapidly changing field components, (2) unshielded operation in the presence of strong radio-frequency interference, and (3) immunity to acceleration-induced forces, resulting in microscopic displacements of components and temperature fluctuations. We believe that the current technology, represented by the SGMS system, which largely uses bulk superconducting components, is reaching its performance limit. More advanced approaches are required to obtain more sensitivity in motion and to provide a technology more practical for fleet deployment.

A technology evaluation and a concept design study were conducted to establish a preliminary concept design and recommendations for development of an advanced all thin-film (niobium) sensor prototype. In contrast to the preceding generation of SGMS system that incorporated bulk niobium technology, sensors using strictly thin-film technology have the potential for increased sensor performance in motion, ease of manufacture, reliability, potential size reduction, and cost reduction. This represents a development analogous to replacing vacuum tube technology with transistor technology.

A contract was awarded to IBM[9] in 1991 to develop an all thin-film gradiometer sensor probe assembly. An artist's conception of this sensor concept, referred to as the Thin-Film Gradiometer, is displayed in Figure 11. A test article has been fabricated to validate the thin-film sensor concept. This test article has provided *the first successful demonstration of a high-sensitivity, all thin-film sensor to operate totally unshielded in a harsh magnetic and electromagnetic environment!* Based on the evaluation of that test article, a final sensor prototype is currently being fabricated to demonstrate a significant increase in performance for naval applications over the older SGMS technology.

Development of an advanced dewar concept was pursued to remove the dewar as a limiting noise source for gradiometer performance. A contract was awarded to Ball Aerospace[11] in 1989 to develop an advanced dewar. To assure this goal, a flexible-design approach, supported by detailed thermal, mechanical, and field calculations, and stringent thermal and magnetic budgets, was established. Fabrication and acceptance testing of the dewar prototype were completed in 1992. *We believe that this dewar provides the most stable passive cooling of a large volume that is available today* (without the use of special complicated approaches, such as feedback-stabilized heating or cooling down to the lambda point of helium). As established from dewar evaluation, it has an isothermal volume of 13,000 cubic centimeters with root-mean-square temperature fluctuations on the order of a micro-Kelvin, sustainable for periods on the order of 100 seconds.

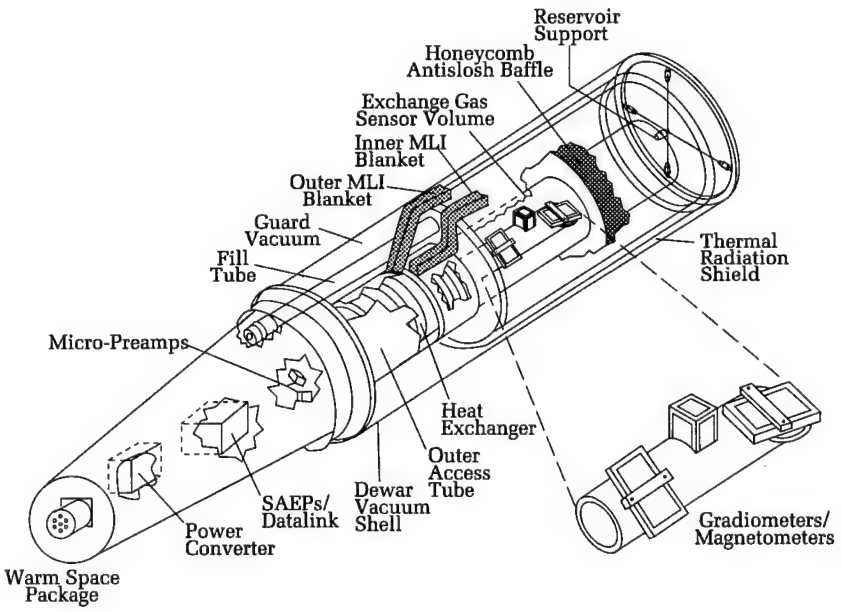


Figure 11. *An artist's conception of the Thin-Film Gradiometer.*

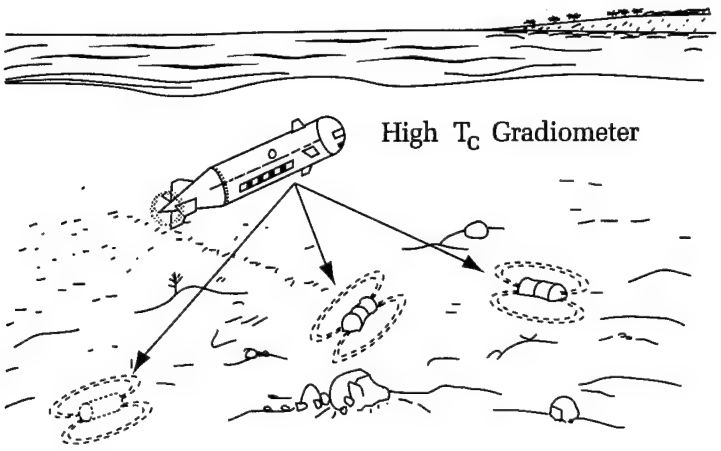


Figure 12. *A multisensor autonomous mine-hunting system concept, including a magnetic detection/classification capability. Integration of a long-range magnetic sensor into an autonomous underwater vehicle is now possible with the advent of high T_c sensors operating in compact nitrogen dewars less than one-half the volume of conventional helium-cooled superconducting gradiometers.*

High T_c Magnetic Sensors and Autonomous Mine-Hunting Systems

High T_c technology has matured (since its inception in 1986) to the extent that practical magnetic sensors are now feasible.[12,13] There has been excellent progress in the development of high T_c Josephson junctions, SQUIDs, and magnetic sensing circuits. In 1993, there were a large number of results reported on magnetometer prototypes. The University of California at Berkeley currently (as of December 1993) holds a record sensitivity of 39 femtotesla per root Hertz (Hz) ($fT/\sqrt{Hz} = 10^{-15}$ tesla per $\sqrt{Hz}$) at 1 Hz.[14]

The development of sensors using the high-T_c materials provides an ease of maintainability and convenience as a result of nitrogen cooling. In particular, a high T_c system with liquid nitrogen refrigeration will offer the opportunity to implement into unmanned underwater vehicle/autonomous underwater vehicle (UUV/AUV) operations (Figure 12) the MCM capabilities already established with the MADOM Project, which involved low T_c technology using liquid helium refrigeration. High T_c sensor packages with acceptable performance are projected to have *less than half the volume of the comparable low T_c sensor*, providing dimensions acceptable for UUV/AUV operation. If liquid nitrogen dewars are used, hold times are projected to be five or six times greater than for comparable helium dewars, which reduces the quantity of cryogen that must be carried on board the ship (in comparison to liquid helium). In addition, using liquid nitrogen reduces supply logistics as a result of the ready availability of nitrogen and simplifies shipboard handling procedures. In comparison to the liquid helium approach, some cost savings are expected in dewar manufacturing and cryogen purchasing.

CSS is actively evaluating this technology and plans in 1994 to initiate development of a high T_c test article using nitrogen-based cooling in order to evaluate one or two high-T_c gradiometer concepts. The high T_c concepts will use more advanced thin-film processing approaches, in comparison to the older bulk technology approaches used for the SGMS system. Hence, these efforts will benefit from the low T_c Thin-Film Gradiometer development described previously.

Other Applications of This Technology

The other most promising naval application area for this technology is antisubmarine warfare (ASW). With the demise of the Soviet Union, the threat offered by their SSBN fleet has subsided. However, there is currently concern for U.S. Fleet security against the proliferation of diesel submarines, which are acoustically quiet and magnetically small. This allows operation in shallow water, which offers a difficult environment for acoustic detection. Improvements in nonacoustic technology, such as those offered by the superconducting gradiometer approaches described here, are needed for airborne search to offset the advantages enjoyed by diesel submarines operating in shallow water.

The introduction of high T_c sensors could open up a range of naval applications as a result of the reduced refrigeration requirements associated with nitrogen cooling. New opportunities include torpedo terminal homing, torpedo/mine fuzing, extremely low-frequency communications, and ASW barriers.

This technology is also an ideal candidate for dual-use applications. The most promising area that embodies the approaches developed for buried-mine detection under MADOM is environmental cleanup. Sensor testing to locate munitions and chemical, biological, and nuclear containers at closing military installations is under serious consideration in coordination with other agencies. Other dual-use opportunities include nuclear regulation, noninvasive biomedical testing, and nondestructive evaluation.

Acknowledgments

The author would like to acknowledge the support from the Office of Naval Research and the Naval Air Systems Command. The author would also like to recognize the leadership of Gary J. KeKelis, who directed this sensor development, and John D. Lathrop and John F. McCormick, who directed the 6.2 and 6.3A MADOM Project. Amongst the many individuals who supported these developments, the author would like to recognize the contributions from David J. Overway, David W. Colberg, Timothy B. Pride, Leonid Vaizer, and Thomas W. Rackers. Finally, the author would like to acknowledge the contributions from industry by Unisys, IBM, Ball Aerospace, and Quantum Magnetics.

References

1. Van Duzer, T., *Principles of Superconductive Devices and Circuits*, Elsevier North Holland, Inc., New York, NY, 1981.
2. Tinkham, M., *Introduction to Superconductivity*, Robert E. Krieger Publishing Company, Inc., Huntington, NY, 1980.
3. Bednorz, J. G. and Müller, K. A., "Possible High Tc Superconductivity in the Ba-La-Cu-O System," *Z. Phys.*, Vol. B64, Sep 1986, p. 189.
4. Wu, M. K., et al., "Superconductivity at 93 K in a New Mixed-phase Y-Ba-Cu-O Compound System at Ambient Pressure," *Phys. Rev. Lett.*, Vol. 58, 1987, pp. 908-910.
5. Josephson, B. D., "Possible New Effects in Superconducting Tunneling," *Phys. Lett.*, Vol. 1, Jul 1962, pp. 251-253.
6. Anderson, P. W. and Rowell, J. M., "Probable Observation of the Josephson Superconducting Tunnel Effect," *Phys. Rev. Lett.*, Vol. 10, Mar 1963, pp. 230-232.
7. Clarke, J., "Advances in SQUID Magnetometers," *IEEE Trans. Electron Devices*, Vol. ED-27, Oct 1980, pp. 1896-1908.
8. The Unisys Corporation, N00612-80-C-8005 (formerly the Sperry Corporation), Defense Products Group, Computer Systems Division, 3333 Pilot Knob Road, St. Paul, MN 55122.
9. IBM Federal Systems Company, N61331-88-C-0105, 9500 Godwin Drive, Manassas, VA 22110, and IBM Research Division, N61331-91-C-0059, P.O. Box 218, Yorktown Heights, NY 10598.
10. Quantum Magnetics, N61331-88-C-0105, 11578 Sorrento Valley Road, San Diego, CA 92121.
11. Ball Aerospace Systems Group, Electro-Optics and Cryogenics Division, N61331-89-C-0072, P.O. Box 1062, Boulder, CO 80306-1062.
12. Clarke, J. and Koch, R. H., "The Impact of High-Temperature Superconductivity on SQUID Magnetometers," *Science*, Vol. 242, Oct 1988, pp. 217-223.
13. Koch, R. H., "SQUIDs Made from High Temperature Superconductors," *Solid State Technology*, May 1990, pp. 255-260.
14. Koelle, D., et al., "High Performance dc SQUID Magnetometers with Single Layer YBa2Cu3O7-x Flux Transformers," submitted to *Appl. Phys. Lett.*

The Author

TED R. CLEM graduated with a B.S. in physics from North Carolina State University, and a Masters and Ph.D. in biomathematics from North Carolina State University, with dissertation research in theoretical statistical mechanics. He joined the research staff at the NSWCCSS in 1982. His research interests include superconducting sensor design and investigation of noise mechanisms associated with superconducting materials. The Chief of Naval Research awarded him for coauthoring the Best Navy Independent Research Paper in 1989. In addition, he received CSS's Award for Special Achievement in Science and Technology in 1985 for his endeavors to understand important noise sources in superconducting sensor technology, and in 1989 for research with the high-temperature superconducting materials. He is also coordinator of low-frequency applications in superconductivity for the Naval Consortium for Superconductivity. He is a member of the American Physical Society. Currently, he manages CSS's Superconducting Gradiometers Project.

Managing the Dynamics of the Electromagnetic Environment to Maximize Combat System Performance

Margaret D. Neel

A major function of the ship self-defense system is coordinating weapons and sensors to eliminate electromagnetic interference (EMI), which otherwise could result in failure to detect and engage targets, loss of missile control, and poor communications. The Naval Surface Warfare Center Dahlgren Division (NSWCDD) has initiated the Combat Systems Frequency Management Program (CSFMP) to manage the dynamics of the electromagnetic (EM) environment (EME) to maximize operational warfighting effectiveness of combat systems. This article shows the value of linking equipment and ship design, acquisition, and life-cycle processes with EM engineering processes involving frequency and other operational restrictions, while accounting for tactical implications. This article also summarizes EMI effects noted in receivers and processors used for weapon control. The dynamics of the operational EME, as well as how these dynamics may be influenced by numerous functions that may not be known or anticipated during a system's development, are discussed.

Introduction

Managing the dynamics of the EME allows combat systems to achieve optimum performance against threats. The manageable dynamic elements include:

- Frequency/Power/Mode
- Time
- Geometry

When EME management fails, resultant EMI degrades specified system performance, often resulting in:

- failure to detect, track, and engage targets
- loss of missiles
- garbled communications
- disruption of flight and other control systems
- receiver damage or burnout

Conversely, effective management of EME dynamics will save lives, ships, and aircraft while also contributing to victory.

In the operational Navy, EMI can manifest itself in numerous ways, as the following examples illustrate. In a 1992 missile exercise, depicted in Figure 1, the scenario included a target at which selected ships could fire missiles (both telemetered and warhead). Ship A, as shown in Figure 1, illuminated the target on a frequency that was not authorized, because that frequency was

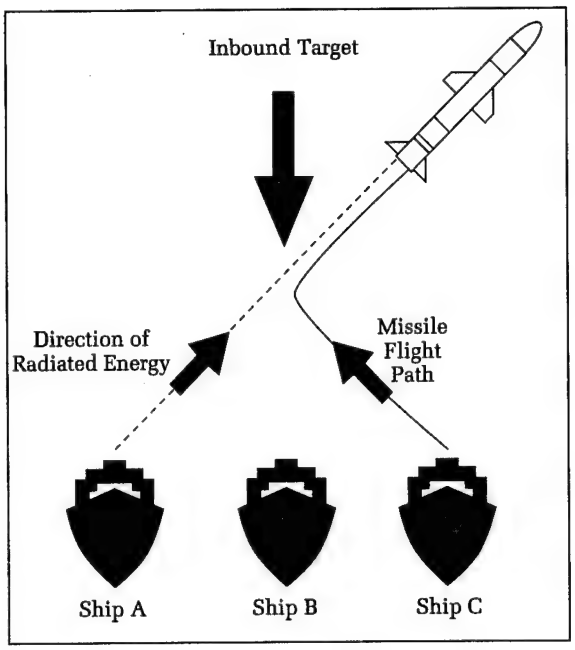


Figure 1. *Missile guidance interference.*

assigned to Ship C. Ship C used its assigned frequency to illuminate the target and launched a missile to guide on that frequency. The target was destroyed by another exercise participant after Ship C's missile launch but prior to the predicted intercept. Ship C then took the appropriate actions to cause her missile to self-destruct. However, the missile detected, received, and guided on the same frequency signal that was erroneously used by Ship A. The missile did not self-destruct, but turned (as shown in Figure 1), continued outbound, and splashed when it reached its maximum range. It is important to note that this incident could have resulted in a serious compromise of safety.

Missile fire-control radars are not the only radars impacted by EMI. Search radars are also seriously impacted, with the worst case being a failure to detect and track targets. In a 1992 North Atlantic Treaty Organization (NATO) exercise, a surface-search radar on a U.S. ship experienced EMI of such severity that the radar was believed to be inoperable.

For given equipment and ship geometry, certain receiver systems are susceptible to burnout from transmitters capable of very high-power operation. To prevent equipment damage, radiation restrictions have been developed to include techniques like sectoring to avoid bearings where harmful interference is probable. Such interference reduces power output, places the system in standby mode and, in the most extreme cases, turns off the system.

In addition to military applications, the EM spectrum is used by private industry, other government agencies, and foreign governments. Treaties and federal laws govern access to the spectrum and each user's priority for access. Overlapping EM spectrum usage between commercial and naval equipment also may result in EMI. Companies representing cellular telephones, satellite communications, bank computers, cable television, and burglar alarms have reported EMI degradation, allegedly caused by U.S. Navy systems. In some reports, EMI to the commercial systems resulted in system shutdown and interrupted service for several hours.

Sharing the spectrum with private industry has imposed additional radiation restrictions on Navy platforms. For example, Navy radars are allowed to operate in the 890 to 902 MHz range on a not-to-interfere basis, due to allocation of these frequencies to commercial land-mobile communications. As a result, the AN/SPS-49 radar, which is currently installed on 135 United States and 16 foreign ships, cannot operate over its entire allocated band. It is limited to 11 of its 48 channels when operating within 200 nmi of the United States and its possessions. Failure to comply with these channel limitations can result in legal action against the Navy and possibly culminate in removal of the offending system from Navy use. However, from the operational Navy's perspective, limiting a system's spectrum use can impact its ability to detect targets in various propagation conditions and can reduce its electronic countermeasures (ECM) capabilities.

Similar restrictions apply in various parts of the world as a result of national interests and treaties. Although some radiation restrictions can be removed during wartime, that authorization is uncommon. In Operation Desert Storm, often referred to as a war, the removal of radiation restrictions was not authorized.

These examples clearly show the need for frequency coordination. Without frequency management, problems such as misidentification of targets, failed missile intercepts, and receiver damage or burnout will continue to plague the Fleet.

Beyond simple spectrum deconfliction, there is a further need to factor tactics and threats into frequency coordination guidance so that EMI control techniques will place minimal restrictions on the Fleet as, for example, an operator's requirement to shift combat system modes to successfully engage threats. The success of our naval operations depends upon a successful interplay among

the operator, tactics, and system engineering. The interactions with other spectrum users from the private sector and foreign governments increase the challenges facing operators, tacticians, and engineers. Near- and long-term resolutions for these challenges involve fusion of the intelligence community, program executive offices, and real-time analysis with the EME management community. This article discusses the role of the CSFMP in meeting these challenges.

Combat Systems Frequency Management Program

Operational Concepts

The Navy Warfare Publication 10-1 Composite Warfare Commander's (CWC) Manual assigns the Electronic Warfare Coordinator responsibility to advise the CWC on all matters pertaining to the use of the EM spectrum. The EM spectrum is subdivided into five "syndicates": radar, electronic warfare (EW), communications, cryptology, and sonar (acoustics). Spectrum management control, however, is decentralized, with each "syndicate" using its own processes.

Within the U.S. Navy, radar frequency management is accomplished through the CSFMP, which is sponsored by the Naval Sea Systems Command (NAVSEA) and executed by NSWCDD. Added resources for expanded Fleet spectrum management are provided by the Space and Naval Warfare Systems Command (SPAWAR), Code 10-12.

Navy Need

In the 1960s, the Navy's "Three T" missile systems—TERRIER, TARTAR, and TALOS—were experiencing operational degradation; too many targets were being missed. A Fleet research missile exercise (FR-69) conducted in the 1960s concluded that EMI affected the TALOS track (fire control) radar and guidance (illuminator) radar. These radars operated in the same frequency band as the TARTAR and TERRIER track radars. Unintended energy was coupling into TALOS radars even when significant frequency separation existed among its operational channels. Application of operational techniques involving frequency separation, platform (geographic) separation, and restricted fields of fire improved TALOS performance. This research exercise demonstrated the need and the effectiveness of spectrum management for improved ship self-defense and overall warfighting capability. Follow-on tests proved the value of expanding spectrum control guidance beyond fire-control systems (FCSs) to include search radars, navigation aids, and carrier-controlled approach radars.

In the early 1970s, CSFMP was initiated to resolve EMI effects to shipboard radars, weapon systems, and navigation aids that result from antenna-to-antenna coupling of signals that share the same frequency band. CSFMP provides this frequency management guidance to the U.S. Navy through operational software and documentation.

CSFMP Products

To support naval spectrum management, CSFMP produces three fleet products:

- NAVSEA T9407-AA-GYD-010 OP-3840, Electromagnetic Compatibility Criteria for Surface Weapon Systems (T9407 OP-3840)
- Electromagnetic Compatibility Analysis Program (EMCAP) software
- "CSFMP News"

T9407 OP-3840. The T9407 OP-3840 publication contains background information on the radars and weapon systems in CSFMP's scope and describes the frequency-management process from the perspective of the operational fleet. T9407 OP-3840 also provides manual procedures for examining the compatibility of frequencies and channels, which may be used if the EMCAP software is not available or if unusual constraints must be accommodated. The EMCAP user's guide is included in T9407 OP-3840.

EMCAP. EMCAP is a DOS-compatible computer program that automates the complicated manual procedures found in T9407 OP-3840. EMCAP was created in the early 1970s when the T9407 OP-3840 publication guidelines became too labor intensive for fleet frequency coordinators. EMCAP provides an automated capability to analyze available radar frequencies and ship formations, and recommends optimal frequency assignments.

"CSFMP News." "CSFMP News" is a semiannual newsletter published by CSFMP. It provides a means of communication among CSFMP engineers and Fleet frequency coordinators, equipment operators, and afloat personnel. The newsletter also includes answers to questions asked by Navy personnel and a "Reader Reply Card" for Fleet requirements and requests for program modifications.

Scope

CSFMP is working toward a goal of frequency management for all naval noncommunications equipment, typically emitters and receivers operating above 400 MHz. This goal is being addressed in incremental phases, by adding more emitters and receivers to CSFMP's scope at each stage. Currently, EMCAP covers the radars and weapon systems shown in Table 1.

Table 1. *EMCAP Scope*

Shipboard Systems

MK 76
- AN/SPG-55B
- CWI

MK 74
- AN/SPG-51C
- OY-88/SPG (Blk 1 & NTU)
- CWI

MK 99
- AN/SPG-62 (CWI)

Basic Point Defense

NATO Sea Sparrow

MK 68
- AN/SPG-53

MK 86
- AN/SPG-60
- AN/SPQ-9
- CWI

MK 92
- CAS
- STIR

3D Radars
- AN/SPS-48
- AN/SPS-52
- AN/SPS-39
- AN/SPY-1

MK 15 Close-In Weapon System

AN/SLQ-32

Other
- AN/SPS-10
- AN/SPS-67
- AN/SPS-55
- AN/SPS-64
- LN-66
- MK 23 TAS
- AN/SPN-41
- AN/SPN-43
- AN/SPN-44
- AN/SPS-40
- AN/SPS-49

Shore-Based Systems

Military ATC (ARSR-1/2/3)
FAA
Cellular Telephones
Cable TV
Satellite Stations
Other Commercial Systems

Aircraft Systems

AN/APS-145
AN/APS-139
AN/APS-138
AN/APS-125
LAMPS (Receivers/ Transmitters)

EMCAP's analysis and EMI resolution techniques address interference that occurs both between emitters on a single platform (intraplatform) and between emitters on different platforms (interplatform). CSFMP's scope of antenna-to-antenna coupling paths includes:

- bistatic (reflected) interference
- sidelobe and backlobe coupling
- intermodulation interference
- known harmonics
- spurious transmissions
- direct radiation

CSFMP Fleet Use

To determine compatible frequency assignments for naval combat systems, EMCAP is executed prior to every naval operation. This includes exercises and operations involving NATO and non-NATO navies, drug interdiction operations with the Coast Guard, training on Naval Reserve ships, combat system ship qualification trials and battle-force deployments. EMCAP is typically executed for the entire group by the lead antiair warfare ship, which then promulgates (via naval message) the frequency assignments to other participating platforms and shore commands (see Figure 2).

CSFMP Life-Cycle Architecture

CSFMP responds to ongoing changes in naval systems. With very rare exception, this response is "back-fit" in nature; that is, systems are addressed only after they reach full operational capability in the Fleet. CSFMP must investigate any changes to naval systems that may affect:

- tactical employment
- operational restrictions
- population
- engineering design
- spectrum usage

The architecture through which CSFMP achieves this responsiveness to Fleet requirements is depicted in Figure 3. The following paragraphs discuss each facet of CSFMP's architecture.

Identify System(s) for Testing

As mentioned above, EMCAP originated as the result of failed missile intercepts in Fleet exercises. The first version of EMCAP addressed only the TERRIER, TARTAR, and TALOS missile systems. T9407 OP-3840 and EMCAP have since expanded to include most—but not all—U.S. Navy shipboard search radars, tracking radars, missile illuminators, and navigation radars (see Table 1). Priority selection of future candidate systems is based upon:

- severity of the EMI
- system priority in mission objectives
- proximity of source and victim systems in Fleet operations
- overlapping frequency ranges
- signal-processing capabilities

Fleet reports of EMI receive top priority in CSFMP's list of radars and weapon systems to investigate. In addition to Fleet feedback, CSFMP engineers monitor various sources of information on ship fit. Through these data sources, CSFMP can respond to new equipment installations, ship alterations, ordnance alterations, and changes in ship status. As ship equipment changes, the probability of EMI can also change. If ship changes involve

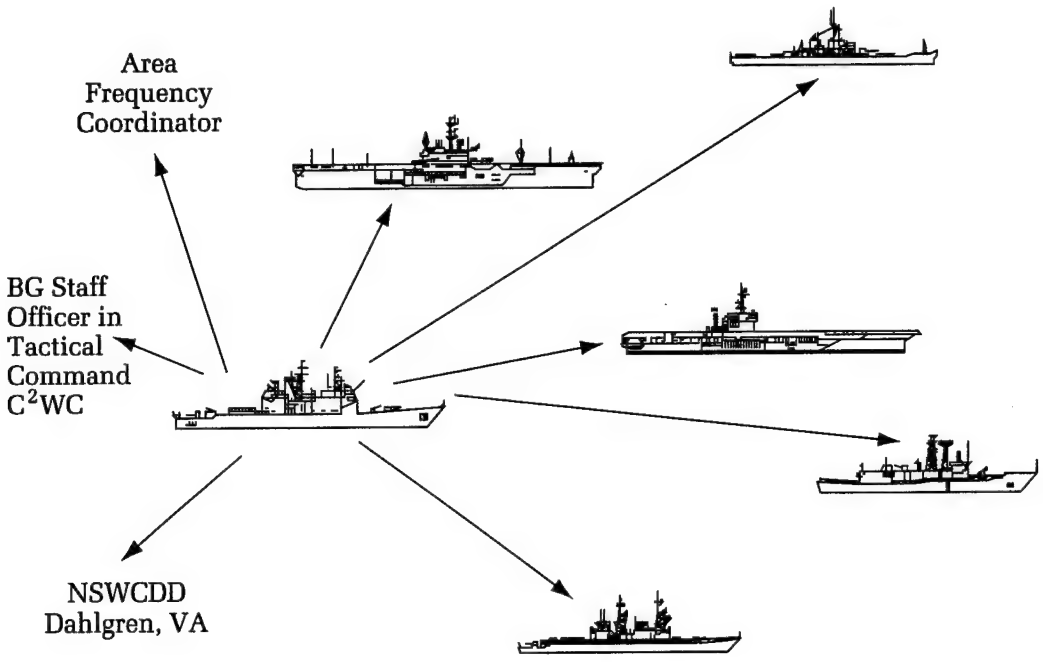


Figure 2. *Fleet use of EMCAP.*

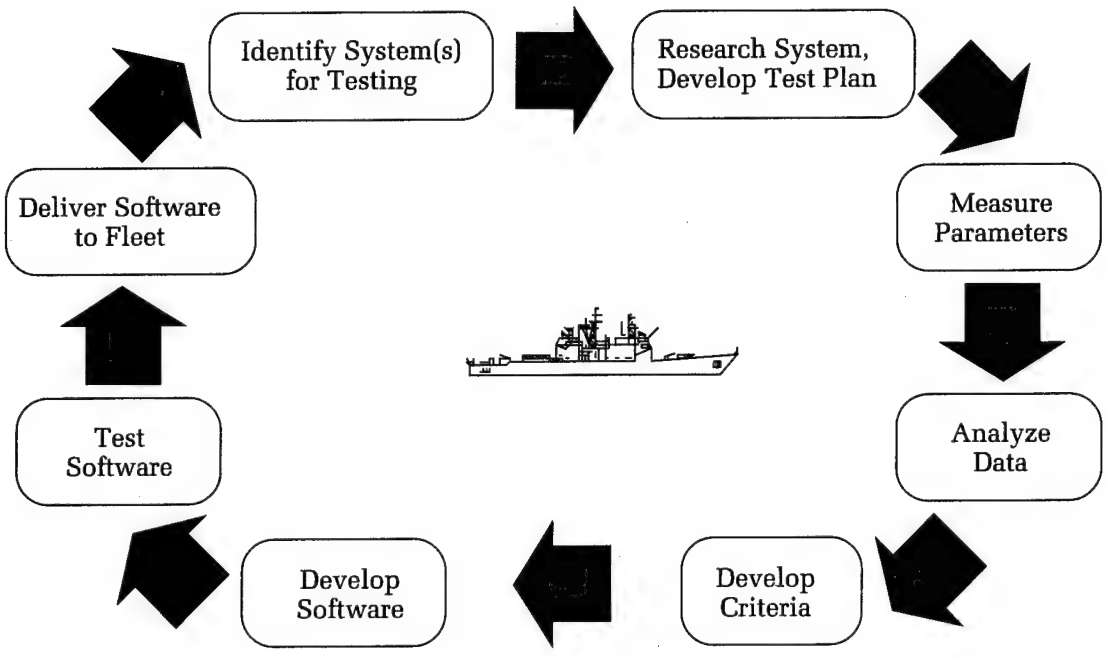


Figure 3. *CSFMP architecture.*

two or more systems in the same or adjacent frequency bands, CSFMP engineers schedule further research into the resulting impact on the EME.

Research System, Develop Test Plan

Once a system has been identified for potential CSFMP tests, the system is thoroughly researched to assess its interference-rejection capabilities, normal Fleet use, population in the Fleet, and other characteristics. From this type of detailed engineering analysis, a test plan is developed. The pretest research helps to ensure that equipment operating modes and operator-selectable options are addressed.

The research also provides a basis for rejecting systems that have been proposed for testing. For example, if the research indicates that the system is scheduled for retirement from the Fleet within a year, expending resources for testing that system is not practical.

Measure Parameters

Based on the test plan, engineers measure the system's parameters and interference-rejection capabilities. The principal parametric data collected include:

- transmitter spectrum
- receiver response
- antenna measurements
- response to other Fleet signals

This data is typically collected in all modes of operation, including the raw or linear video mode, to allow for maximum operator flexibility in naval operations.

Historically, system interference rejection measurements were conducted by scheduling two ships for dedicated tests. The ships would vary their separation distance and the system's frequency and operating modes, while EMI engineers monitored the EMI victim's ability to detect or track a small target. The implications of this test method include the following:

- The system cannot be tested until after it has been installed on Navy platforms.
- Scheduling was often difficult and fluid.
- The overall cost in time and materials was high.
- Yet, the confidence in the measurements was high.

To minimize the negative factors listed above, CSFMP has developed a hardware suite that uses analog-to-digital and digital-to-analog converters to record and play back the system transmission. CSFMP's digital recording suite is portable and can be used when only one radar system is available for measurement. Confidence in measured results remains high, while scheduling, availability, and cost factors are minimized.

Over CSFMP's 20-year history, numerous approaches have been investigated for obtaining the necessary system data. Various models and specification data have been tested but did not achieve the accuracy, reliability, and cost effectiveness obtained by using the aforementioned digital recording and playback method.

Analyze Data

Upon completing the measurements, the data is analyzed to determine guidelines to minimize EMI. The guidelines are based upon the following parameters:

- sensitivity of the interference victim radar to the source radar waveform
- power of the interference source radar
- antenna pattern of the source radar antenna
- antenna pattern of the victim radar antenna
- basic transmission loss
- off-frequency rejection in the victim radar

From these parameters, the amount of EMI a system will tolerate in selected modes of operation can be calculated. The data analysis is becoming increasingly complex due to factors such as the commercial use of frequencies and joint and combined military operations.

Develop Criteria

The Electromagnetic Compatibility (EMC) criteria produced by CSFMP engineers is typically in the format of frequency/distance tradeoff tables. An example is shown in Table 2. When systems that operate in the same frequency band are stationed closer

Table 2. *Frequency/Distance Tradeoff*

Distance		Frequency		Total Loss (dB)
Separation (nmi)	Loss (dB)	Offset (MHz)	Loss (dB)	
70	215	0	0	215
43	200	20	15	215
32	185	30	30	215
30	178	50	37	215
28	170	300	45	215

together, they must maintain a greater frequency separation in order to adequately filter out each other's signals. The smaller the separation distance, the larger the required frequency separation for compatible operation. Based on previously measured parameters, the example system in Table 2 can achieve its intended performance as long as interfering signals are 215 dB down. In other words, this example system will tolerate potentially interfering signals if the two systems are separated by 70 nmi even if the systems are operating on the same frequency. The two systems are compatible at a closer distance of 28 nmi if a 300-MHz frequency separation is employed.

The criteria developed at this stage include general guidance on system spectrum use and the step-by-step manual procedures for frequency coordination. The criteria also address any assumptions or conditions, such as sea-state or propagation ducts, which affect real-world operations.

Develop Software

Using a variety of techniques, including heuristics, artificial intelligence, and operations research, the EMC criteria are translated into software for EMCAP. The software expands upon the manual coordination procedures in T9407 OP-3840 and implements techniques that cannot be manually duplicated in a timely manner. The software also adheres to CSFMP standards for structure, design, portability, maintainability, and speed of execution. The software algorithms include aspects of the tactical employment of the system to the greatest possible extent. System capabilities and intended function (e.g., self-defense versus long-range defense) influence software prioritization in allocating limited spectrum resources. The technique of "graceful degradation" is factored into all software assignment routines. Graceful degradation is necessary when the number of frequency assignments required exceeds the number of available frequencies. This imbalance between frequency supply and demand is common in most naval forces of eight or more ships, and is exacerbated in littoral areas due to close-ship spacing and shore-based emitters.

Test Software

The software is extensively tested prior to fleet delivery. Testing is conducted on a wide variety of potential host computers, from the least capable to the most capable configurations. The feasibility of the EMCAP assignments and operational work-arounds in Fleet operations is examined and verified by Navy personnel from tactics and training commands. Although it is not possible to test all possible combinations of ships (683 factorial, based on the current EMCAP scope of 365 U.S. and 318 foreign ships) and all possible combinations of frequency assignments, the software integrity is determined through both execution and code analysis. Since 1985, when CSFMP began keeping records on software performance, EMCAP has never abnormally terminated in Fleet use.

Deliver Software to the Fleet

Fleet delivery of EMCAP and the associated document, T9407 OP-3840, occurs on a 12- to 18-month cycle. T9407 OP-3840 is distributed in printed format; EMCAP is distributed on 3.5- and 5.25-inch floppy disks for execution on DOS-compatible personal computers.

In the U.S. Navy, T9407 OP-3840 and EMCAP are currently available on board 365 ships of the U.S. Navy and Coast Guard to support fleet operations. U.S. Navy and Coast Guard distribution is listed in Table 3.

Table 3. *T9407/EMCAP Distribution*

- Cruisers (CG/CGN)
- Carriers (CV/CVN)
- Destroyers (DD/DDG)
- Frigates (FF/FFG)
- Command Ships (LCC/AGF)
- Assault Ships (LHA/LHD/LPH)
- Oilers (AO/AOR)
- Fast Combat Support Ships (AOE)
- Mine Countermeasures Vessels (MCM)
- Destroyer Tenders (AD)
- Submarine Tenders (AS)
- Ammunition Ships (AE)
- Cargo Ships (LKA)
- Transport Dock (LPD)
- Landing Ship Dock and Tank (LSD/LST)
- Combat Stores (AFS)
- Repair Ships (AR)
- Salvage Ships (ARS)
- Patrol Hydrofoils (PHM)
- Cruiser and Destroyer Groups and Squadrons
- Coast Guard Cutters (WHEC, WMEC)

Table 4. *CSFMP T9407/EMCAP Data Base*

365 Platforms (USA)

- Aircraft Carriers
  CV/CVN
- Cruisers
  CG/CGN
- Destroyers
  DD/DDG
- Frigates
  FF/FFG
  FFT
- Mine Warfare
  MCM
- Amphibious
  LCC LHA LHD LKA
  LPD LPH LSD LST
- Auxiliary
  AD AE AFS AGF AO
  AOE AOR AR ARS AS AVT
- PHM Patrol Hydrofoils
  PHM
- Military Sealift Command
  TAE TAH TAGFF
- Coast Guard
  WHEC WMEC
- Aircraft
  LAMPS
  E-2C Hawkeye

318 Foreign Platforms (NATO and Non-NATO Countries)

Australia	Belgium	Brazil	Canada
Chile	Colombia	Denmark	France
Germany	Greece	Italy	Japan
Korea	Mexico	Netherlands	Norway
Pakistan	Peru	Portugal	Saudi Arabia
Spain	Taiwan	Thailand	Turkey
Tunisia	United Kingdom	Vietnam	

Although T9407 OP-3840 and EMCAP include 318 foreign ships in their data base, neither product is releasable to foreign countries in their current form. Table 4 lists the U.S. ship types and the foreign countries represented in CSFMP's T9407 OP-3840 and EMCAP data base.

Ongoing Fleet Support

Upon request, CSFMP personnel provide the Fleet with hands-on training for EMCAP execution. The training occurs aboard ship, at schools, and at predeployment meetings. Historically, the radar spectrum management training was primarily held aboard ship. During the last two years, the majority of this training has transitioned to schools, such as the Atlantic and Pacific Tactical Training Groups, the Naval War College, and the AEGIS Training Center.

CSFMP has also supported virtually every U.S. operation, including those involving NATO and non-NATO countries. Upon request, CSFMP engineers have developed frequency plans to accommodate unusual constraints, close ship stations, threat-detection frequencies, and foreign emitters not included in EMCAP. Through the exercise support and the training, a wealth of data has been collected regarding the capabilities and limitations of T9407 OP-3840 and EMCAP, both in terms of engineering criteria and operational employment. The limitations and Fleet recommendations for improvement are incorporated each fiscal year into CSFMP's development architecture.

The Need For Change

In today's world, the EM spectrum is becoming more congested and politically charged. U.S. Naval strategy is moving toward regional and littoral conflicts. Future operations will involve joint U.S. forces combined with allied units, with their multitude

of EM equipment vying for use of the available spectrum. In addition, both domestic and foreign industry want part of the available spectrum in order to pursue profit.

These new environmental and industrial issues, as well as Fleet upgrades and modernization, are affecting the original needs and requirements for EMCAP. Factors that impact spectrum management and that must be considered are:

- new and modified systems
- battle-force size
- naval aircraft, joint, and multinational emitters
- shore proximity
- commercial emitters and receivers
- electronic warfare
- tactical and real-time considerations

A better understanding of self-defense systems' performance against threats, as well as changes and improvements to self-defense system capabilities, evolving threats, and a growing number of emitters, have thrust frequency-management into a prominent role in self-defense performance. These factors are further highlighting the importance of frequency management as a tactical tool. In order to accommodate the dynamic nature of tactics, frequency-management analysis and EMI resolution must become available in real time to the operational Fleet.

Battle proven, EMCAP performed well during Desert Storm, but deficiencies were noted. Operations Desert Shield and Desert Storm clearly demonstrated how the "reactive back-fit" process affects the fleet. During the Gulf War, the MK 74 FCS was undergoing a new threat upgrade (NTU). This upgrade changed the spectrum utilization of the MK 74 fire-control radar. As MK 74 NTU ships joined the coalition forces in the Persian Gulf region, incidents of EMI increased in the 4000 to 8000 MHz Radar Band (C-band or Institute of Electrical and Electronics Engineers 4000 to 6000 MHz Frequency Band known as G-band).

CSFMP engineers responded during Desert Shield by developing written procedures to manually allocate portions of the spectrum to the MK 74 NTU radar. Naval message traffic reported that the procedures were effective, well written, and easy to follow; however, frequency coordination became a full-time job at a time when manpower was critical. As Desert Shield transitioned to Desert Storm, CSFMP completed modifications to EMCAP to automate the MK 74 NTU coordination procedures. The improved version of EMCAP was immediately distributed to all U.S. Navy combatants. Although CSFMP was able to respond to Fleet needs during the war, the timeliness could have been improved if the MK 74 NTU track radar had been tested for frequency compatibility with other C-band systems prior to the NTU deployment.

CSFMP's architecture is currently back-fit in nature. EMI resolution techniques are typically adopted after the equipment has been designed and fielded. Unfortunately, these back-fit techniques invariably limit or curtail system capabilities. The impact of the EME should be assessed and addressed when the system is in its engineering design phase.

New and modified systems need to be added to EMCAP during their development stages ("forward-fitting" versus "back-fitting"). In many cases, CSFMP is not aware of system development and deployment and, therefore, cannot perform EMC measurements until after the system has been fielded. CSFMP's primary source of information on interference comes from the afloat Navy in the form of missile firing reports. With CSFMP's average upgrade cycle of 15 months, responding to EMI effects only after it has adversely affected missile performance in naval operations does not adequately support our nation's self-defense and warfighting capabilities.

Desert Storm taught other lessons. EMCAP, as originally conceived, could coordinate only up to 15 ships in one group. This limit was acceptable, because task forces were not expected to contain a larger number of missile ships in the 1970s. Desert Storm dispelled that notion. Well over 100 ships from several nations were involved.

Due to the battle group's increase in size, the quantity of systems needing EMCAP coordination has changed. EMCAP originally provided spectrum management for surface ship search and fire-control radars and illuminators. The target population now includes: naval aircraft radars, missile seekers, identify friend or foe, joint-forces emitters, land-based emitters, multinational-military emitters, and EW systems.

To accommodate these systems, the new geopolitical environment, and today's sophisticated threats, EMCAP needs to provide tactical, real-time spectrum management. Current frequency assignments are established by the EMCAP frequency coordinator prior to deployment. Generally, follow-up actions are minimal. But now, different threats, more emitters, upgraded system capabilities, and a better understanding of system performance against threats is driving frequency management to become an active option in tactics.

Furthermore, radar management must be fused with the other four "syndicates" (EW, communications, cryptology, and sonar) to optimize all the dynamics of the EME to maximize the performance of ships' self-defense systems.

CSFMP Status

Meeting the requirement for real-time frequency management is a challenge that CSFMP has recently undertaken. Efforts to develop real-time spectrum management have begun. A proposal for an Advanced Technology Demonstration (ATD) has been drafted and approved for FY 96 funding; however, much work remains.

Although EMCAP can be expanded to provide naval air and multinational-military frequency management, this increased scope requires an innovative approach. Rather than basing frequency separation solutions on empirical measurements, the solutions must be based on analytical tools, which do not have the accuracy of empirical measurements. Much work has been done on management issues related to proximity to shore and conflicts with commercial equipment. This work must continue and expand to adequately address ship self-defense in a littoral environment. Only a small amount of work has been accomplished regarding tactical issues in frequency management. Self-defense system tactics for hard kill and soft kill must be integrated to ensure spectrum coordination and to reduce fratricide.

Proposed Program Process

The proposed improvements to the program process are divided between near- and long-term objectives.

Near-Term Frequency Management Process

CSFMP's near-term objectives for process improvement include the following goals:

- Require the establishment of Frequency Management Advisory Boards. These boards should be composed of the Program Executive Office's (PEO's) system engineers, CSFMP engineers, and Fleet operational personnel. The charter for the Board shall be to optimize frequency-management criteria for maximum efficiency of the combat system in the context of threat hard kill/soft kill in the operational EME. Rules relating to performance against specific threats should be established as guidance for dissemination and use by the Fleet.
- Establish the following new equipment acquisition milestones for PEOs:
  - Frequency Management Advisory Board for EMCAP convened
  - EMCAP tests scheduled
  - EMCAP test plan developed
  - EMCAP tests complete
  - EMCAP data analysis complete
  - EMCAP management criteria developed
  - EMCAP software complete
  - EMCAP software testing complete
  - EMCAP software delivered to the Fleet
- Shift primary funding responsibilities to PEOs or other acquisition managers for EMCAP modifications when new equipment is provided to the Fleet, or when the performance of existing equipment in the Fleet is modified.

These proposed short-term improvements to the CSFMP process are designed to bring together the PEO system engineers, CSFMP criteria development engineers, and tacticians (e.g., Surface Warfare Development Group). These changes establish links among the activities that guide naval use of hard-kill and soft-kill systems. This will ensure that:

- optimum frequency-management criteria are provided to the Fleet for effective system utilization for ship self-defense,
- all participants agree to and have inputs to the frequency-management criteria, and
- the most effective rules are established for frequency management relative to system effectiveness, especially for hard-kill/soft-kill optimization.

Long-Term Frequency Management Process

Long-term objectives, as defined in the ATD proposal, involve shifting from a static to a dynamic on-line radar frequency management program and then fusing this improved program with the other four "syndicate" management groups (EW, communications, cryptology, and sonar) to form a new program for centralized EM management under the control of the Navy's Command and Control Warfare Commander (C^2WC). Figure 4 depicts the dynamic, on-line spectrum management system as proposed in the ATD.

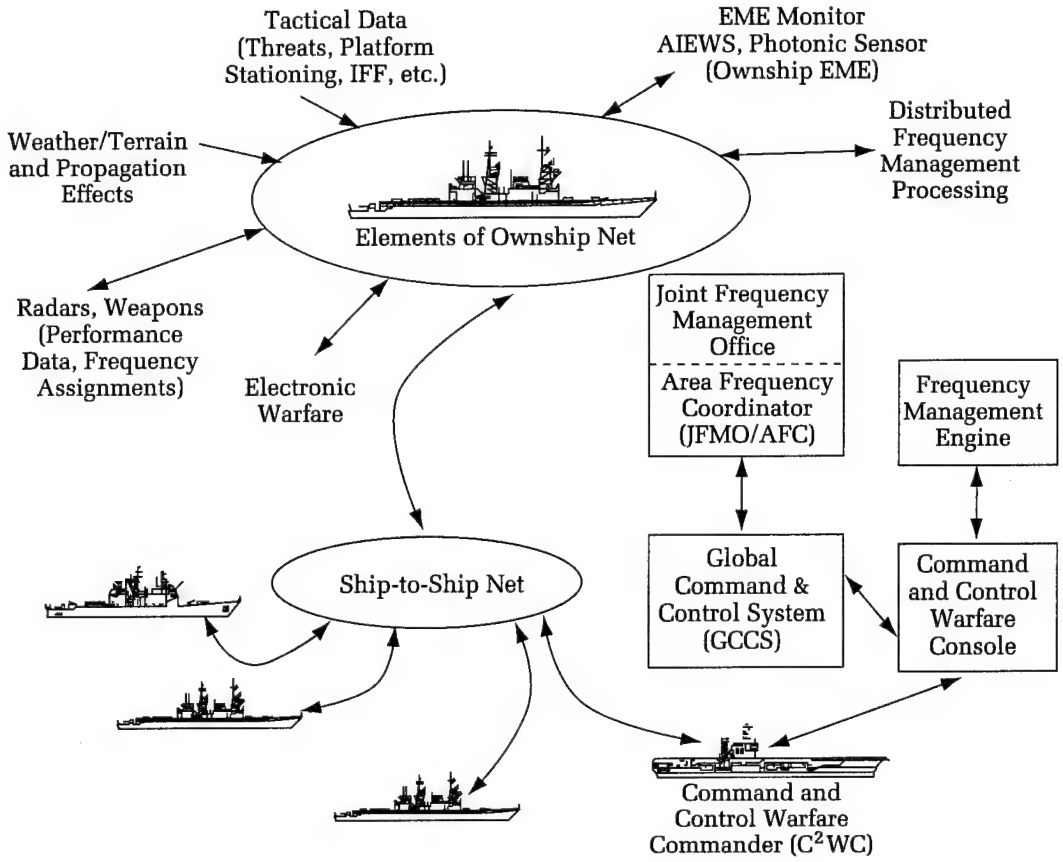


Figure 4. *Dynamic, on-line spectrum management system.*

The long-range goal includes the integration of current frequency management software with intraship self-defense systems using artificial intelligence and ship-to-ship data links similar to the Integrated Interior Command and Control [(IC)^2], Cooperative Engagement Capability (CEC), or LINK 16 Joint Tactical Distribution System (JTIDS). Emissions from each ship's self-defense systems will be coordinated for minimum intership EMI and maximum effectiveness. Sensors, such as the Advanced Integrated Electronic Warfare System (AIEWS) or a sensor from evolving photonics technology, will monitor the EME and identify radiated frequencies to the accuracy necessary for spectrum coordination (approximately 1 MHz for systems above 400 MHz). This monitored emissions data for each ship will be provided on a network for improving hard-kill and soft-kill effectiveness; emissions control; interoperability with joint, combined, and civilian systems; and other applications. An interactive interface will be provided to each ship self-defense system for the transmission of frequency-spectrum commands. The commands will maximize own-force hard-kill and soft-kill self-defense effectiveness by changing intership emitter frequencies and other characteristics as needed to achieve coordinated use of the EM spectrum.

It is envisioned that EMCAP will be enhanced and installed on a shipboard tactical workstation computer as the baseline for spectrum management command and control (C^2). Using frequency-management criteria and tactical rules for littoral and coalition engagements, artificial intelligence software will be developed for control of own-force emissions, with maximum tactical flexibility and a minimum of hard-kill/soft-kill fratricide. The future EMCAP C^2 software will provide required interfaces to EW, Signals Intelligence (SIGINT), communication links, and Ship Self-Defense Systems (SSDS) to permit effective interactive intership operation. The interface to EW (e.g., AIEWS), SIGINT, and SSDS will permit a coordinated handshake with these functions and coordinate the "syndicates" of the EM spectrum. The interface with Electronic Support Measures (ESM) and ECM systems will allow interactive adjustments in own-force emitters to maximize ESM detections and/or reaction capabilities.

Sensors will be developed to monitor each ship's emitters, with the information provided on a data link. Where improper emitter characteristics exist for compatible intership operation, the EMCAP C^2 software will provide on-line, real-time commands for changing the conflicting emitter's output. Monitored real-time data will provide emission status and characteristics of own-force emitters for use in EW/SIGINT filtering or SSDS processing. A function such as ESM could request momentary "emission free" windows within the spectrum, and the EMCAP C^2 software would appropriately command retuning of own-force emitters. When ECM (jamming) is to be employed, the software could command retuning of hard-kill emitters and receivers to minimize jamming effects on own-force systems.

Through the integration of ship-positioning and mapping data, the EMCAP C^2 software will provide the C^2WC with automated, on-line, dynamic control of ship-to-ship or ship-to-shore radiation restrictions. As ships move closer to each other or to the shore, the EMCAP C^2 software will assess own-force emitters' spectrum use and generate commands to retune emitters as needed to (a) ensure interoperability between ships' sensors and (b) prevent harmful EMI to land-based systems or forces ashore.

Capabilities of this future integrated system follow:

- Provide a dynamic, quick-reaction spectrum management capability with on-line monitoring
- Provide a networking capability to obtain necessary data bases and environmental data, and to disseminate spectrum use information
- Provide a frequency management plan that considers:
  - environmental and propagation conditions
  - other service assignments
  - platform warfare duties
  - system capabilities, limitations, and tactics
  - threat characteristics
- Provide operational degradation data for each system, platform, and the overall force and identify potential vulnerabilities
- Compute platform-to-shore restrictions
- Graphically drive the software input/output (I/O)
- Provide network interference reporting capability
- Provide an analysis tool to rapidly deconflict interference

Navy Benefits

Adopting the process proposed in this article will allow the Navy to modify the operational dynamics of the EME to correct for EMI that is already seriously affecting system performance. The proposed dynamic, on-line spectrum management system will integrate spectrum use of U.S. Navy combat and EW systems, provide a readiness assessment of self-defense systems, and optimize the effectiveness of ship self-defense hard-kill and soft-kill systems.

The Author

MARGARET D. NEEL is the project manager for EMCAP at NSWCDD in Dahlgren, Virginia. She has 10 years of experience analyzing, designing, and developing radar signal analysis software. Her research includes artificial intelligence, heuristics, and operations research techniques. She has supported numerous Naval deployments and is recognized in the Fleet as a leading authority on naval radar spectrum utilization. Ms. Neel received Bachelor of Science degrees in computer science in 1982 and mathematics in 1983 from the University of Kentucky.

Electronic Warfare in Ship Defense

Thomas W. Kimbrell

In response to the increasing sophistication of the threat, the complexities of the littoral environment, and fiscal constraints, our electronic warfare (EW) systems must make increased contributions to ship self-defense. This article addresses the threat and environment that future EW systems must face, the role EW must play in the combat system, and the improvements in self-defense that an integrated EW capability brings. It also defines those parameters that are key to achieving the performance required of our next-generation EW systems and identifies the requirements that drive these performance parameters.

Introduction

EW encompasses a broad range of functions. It includes actions both to exploit the enemies' use of the electromagnetic spectrum and to preserve the friendly use of that spectrum. EW is defined as military action involving the use of electromagnetic energy to determine, exploit, reduce, or prevent hostile use of the electromagnetic spectrum and action that maintains the friendly use of the electromagnetic spectrum.[1] Traditionally, EW has been broken down into signal intelligence (SIGINT), electromagnetic support measures (ESM), electronic countermeasures (ECM), and electronic counter-countermeasures (ECCM). A more complete taxonomy is given in Figure 1. Both SIGINT and ESM encompass the interception, analysis, identification, and location of hostile or potentially hostile emissions. SIGINT is non-time-critical and strategic in nature; whereas ESM is tactical in nature, though at times the distinction is blurred, as the tactical use of information from primarily electronic intelligence systems becomes more common. The distinction between communications intelligence and communications ESM is also hazy, because cryptologic systems with direction-finding capabilities can provide tactically significant information. Although there is a trend across the entire EW spectrum toward more integration of EW into our combat systems, this article will focus on the areas highlighted in Figure 1, which are sometimes referred to as tactical EW.

The role and integration philosophy of naval tactical EW systems are undergoing fundamental changes in response to the changing radio frequency (RF) environment and evolving threat to surface combatants. EW has played a significant role in naval warfare since World War II, when high-frequency communications and radar-intercept receivers played significant roles in the U-boat war in the Atlantic.[2,3] The stand-alone design philosophy of currently deployed EW systems can be traced to the signal environment that existed from World War II through the 1970s. During that period, the RF signals to be intercepted were limited in complexity and numbers, and the required system reaction times were much less demanding than they are today. As shown in Table 1, the RF environment has increased in complexity and density. The

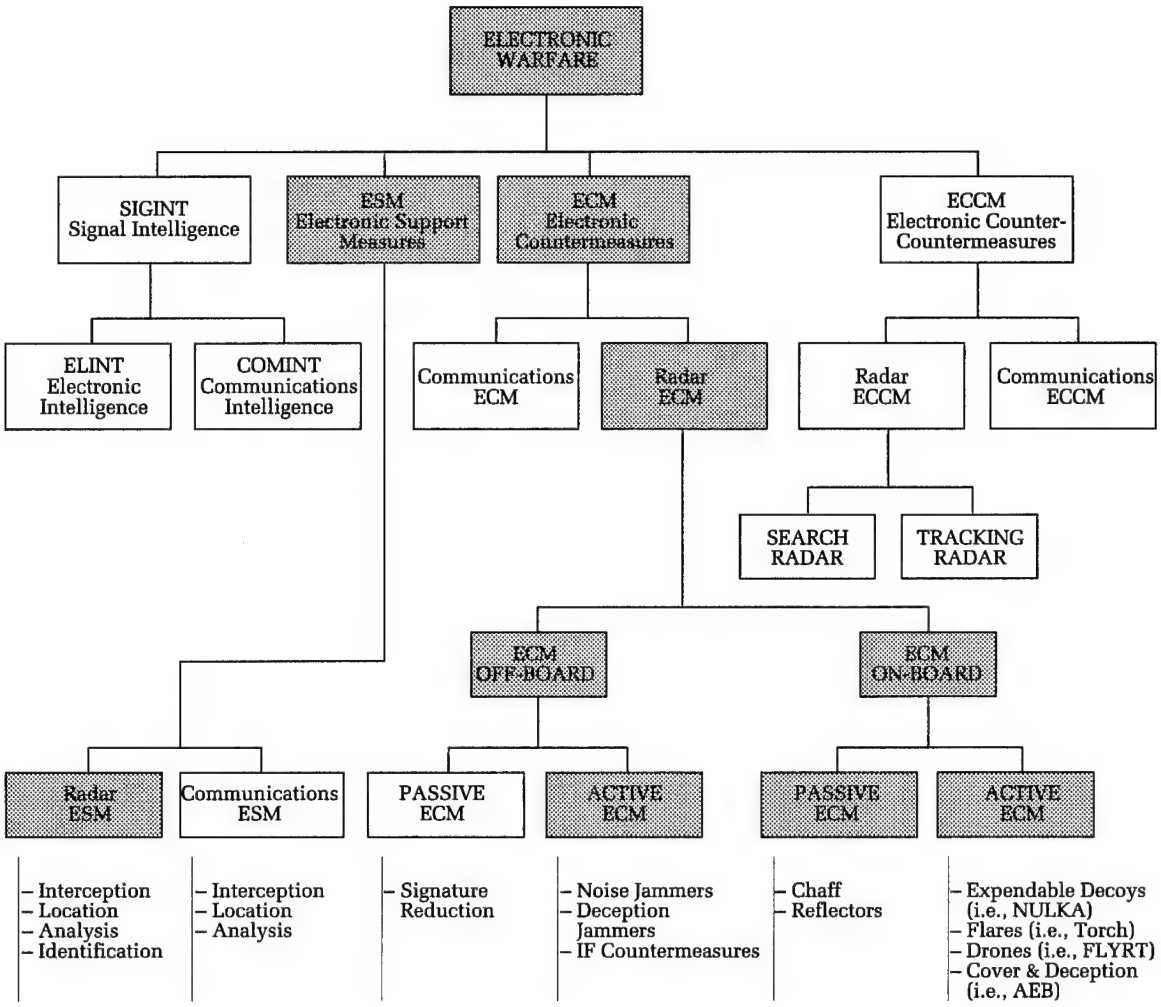


Figure 1. *EW taxonomy.*

introduction of the antishipping guided missile presented naval tactical EW with a new mission. The AN/SLQ-32 (Figure 2) was the Navy's EW response to that threat. With the AN/SLQ-32, naval shipboard EW shifted from a surveillance and situational awareness emphasis to self-defense, and began to be considered primarily an antiair warfare (AAW) asset. EW also continued its stand-alone design philosophy. That design philosophy is no longer adequate to handle the threats and RF environment of the future.

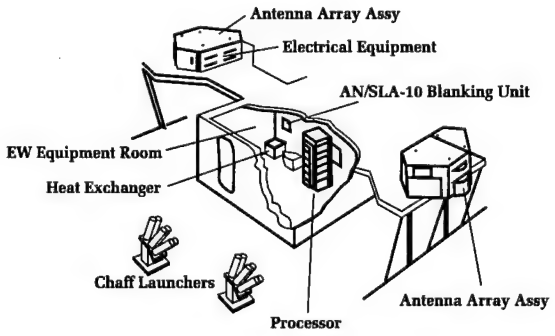


Figure 2. *AN/SLQ-32.*

EW Environment and Threat Trend

The EW environment and threats our EW systems must detect and counter are rapidly changing. Table 1 shows the increase in pulse density and waveform complexity between the 1970s and 1990s. An ESM system not only has to detect and track pulses emitted from threats, but also has to sort out those pulses from the hundreds of thousands to millions of pulses and continuous wave emissions from neutrals and friends. This must often be done in severe weather and propagation conditions, and in the presence of hostile ECM. EW systems must meet their performance requirements in these severe environments.

Table 1. *Radar Environment*

CALENDAR DECADE	1970	1980	1990
Maximum Signal Density	40,000 pps	500,000 to 2,000,000 pps	1,000,000 to 10,000,000 pps
Frequency Range	Selected portions of 2–12 GHz	Selected portions through 40 GHz	Selected portions through 40 GHz
Pulse Repetition Interval (PRI)	Stable, multiple-pulse trains	Stable, jittered, staggered	Stable, jittered, staggered, pseudorandom
Radar Systems Features	Single frequency	• Multiple frequencies • Frequency hopping	• Multiple frequencies • Frequency hopping • Spread spectrum • Multiple agile antenna beams
	Interpulse processing	• Intrapulse phase shift • Coded modulations • Power management • Digital processing • Increased duty cycles with lower peak power • Weapon systems using dual-mode seekers (IR, RF)	• Intrapulse phase shift • Coded modulations • Power management • Digital processing • Larger time bandwidth product radar signals • Increased duty cycles with lower peak power • Weapon systems using multimode seekers (IR, laser, active, and passive RF)

The task at hand is significantly easier in the open ocean than it is in today's littoral environment. In the open ocean, few neutral emitters would be close enough to be detected, and the general direction of potential threat emitters would be known. In the littoral environment, most of the pulses received will be from friendly or neutral or, possibly, quasi-neutral emitters. These signals will include both commercial and military emitters, and will come from the land, as well as the sea and air. Ownship and nearby emitters, and reflections of these emitters off ships, land, and maritime platforms, can be expected to contribute significantly to the signal environment.

In this rapidly changing environment, there are several emerging threat-emitter trends. Military radars are becoming more agile in PRI and frequency, and incorporate more ECCM features. Increased numbers of pulse-doppler and high-duty emitters severely complicate pulse-sorting requirements. As older missiles, and targeting and surveillance radars are retired, more and more threats use monopulse tracking. There is an increased emphasis on low probability of intercept through the use of power-managed waveforms, spread-spectrum techniques, noise-like waveforms, and frequency agility over a large frequency extent. As targeting and inertial navigation systems have become more sophisticated, antishipping missiles (ASM) are using late emitter turn-on to reduce EW system reaction time. They often use narrow pulse widths that make chaff less effective, while providing sufficient energy for detecting and tracking at the shorter required ranges. With the advent of solid-state radars for surveillance, targeting, and fire control, there is an opposing trend toward much longer pulse widths, with intrapulse coding used to preserve range resolution. These solid-state radars are generally wider in frequency bandwidth than radars using tubes and are inherently waveform diverse. They are also higher in duty factor and lower in peak power. These features combine to make pulse trains more difficult to detect and de-interleave into individual

pulse trains. The higher duty increases the probability of pulse overlap and ties up circuitry used to measure pulse width.

Electronic Warfare Systems Role in the Combat System

In response to the increasing sophistication of the threat, the complexities of the littoral environment, and fiscal constraints that require us to achieve as much synergism among our various weapon systems and sensors as possible, the role of tactical EW in our combat systems is changing. EW must provide a layered self-defense capability operating in parallel and in coordination with our missile and gun systems. It must also function as a combat-system sensor, providing situational awareness, emitter and platform identification, and long-range target tracking. Likewise, other combat-system sensors must be considered valuable sources of information for our EW systems. Today's digital technology makes this required sharing of data among all of our combat system elements feasible. In addition to supporting AAW, today's EW systems must support all warfare mission task areas, including Anti-Surface Warfare, Strike, and Space and EW, in both littoral and open-sea environments. Given sufficient sensitivity, modern ESM systems can provide excellent over-the-horizon targeting capabilities, especially if systems on multiple platforms are used to localize targets.

In today's political reality, taking even a single hit is unacceptable. A reasonable goal might be for a ship to have a 95-percent probability of being able to sustain either five attacks or a single attack of five threats. The probability of being able to sustain N attacks (assuming independence and equal probability of each engagement) can be computed from Equation (1) and is plotted in Figure 3.

$$\text{Probability of Success} = 1 - (1 - P_{SE})^{\text{Number of Engagements}} \quad (1)$$

As shown in Figure 3, no single system will be able to provide the desired effectiveness over many engagements. The cost of building a 99-percent effective system needed to provide a 95-percent probability of surviving five attacks (or a single attack by five threats) is probably prohibitive even if the technology to build such a system were available. This level of robust performance can best be achieved by increasing the number of engagement opportunities. This can be accomplished by engaging with multiple systems, by engaging the threat multiple times with a given system (layered defense), or by a combination of both.

As shown in Figure 4, EW systems have many potential opportunities to engage the threat and increase overall self-defense effectiveness. In the past, ECM has focused on the end game of the ASM engagement, attempting to break the range or angle track loop of the missile guidance system with active ECM and seduction chaff, or attempting to prevent missile lock-on with distraction chaff. Breaking a locked-on track loop is a difficult challenge to ECM, especially if the tracking servos have settled and are operating in a narrow-band mode. Break-lock requires high radiated power and large chaff

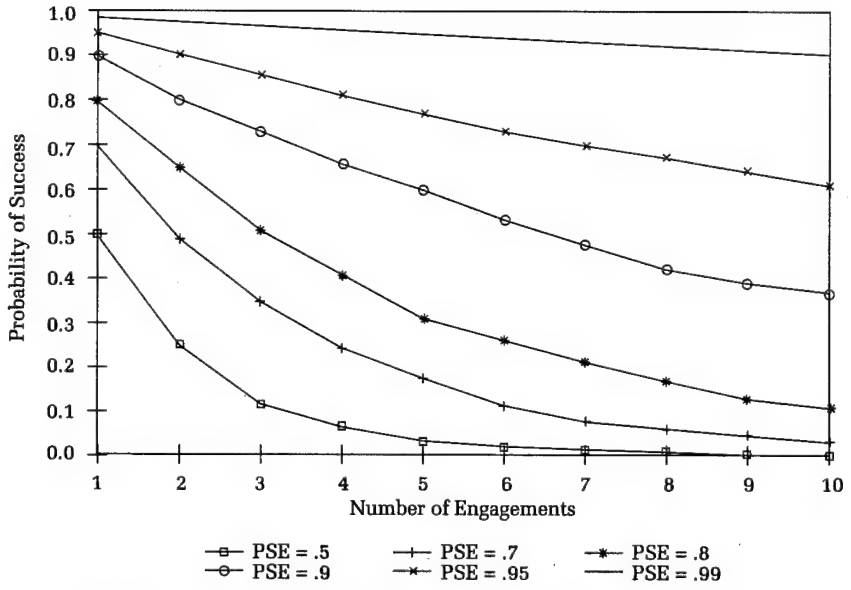
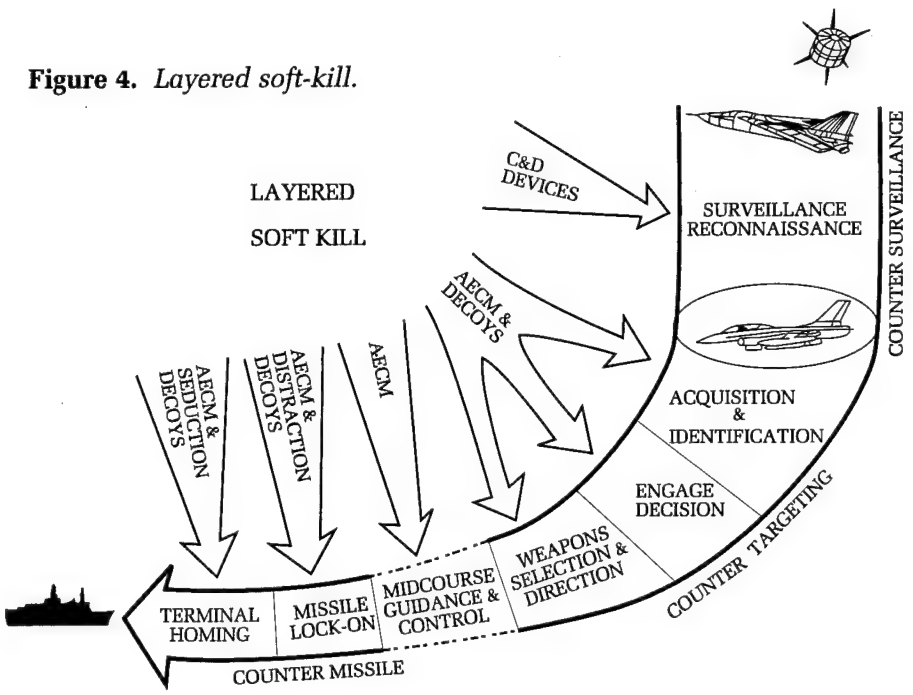


Figure 3. *Probability of being successful for all engagements.*

radar cross section and often requires detailed knowledge of the threat to be effective. The active ECM waveforms or chaff return must compete against the return of the ship being defended.

By combining layered EW engagements with missile and gun systems engagements, it is feasible to reach the desired level of effectiveness with achievable, single-system performance levels. Figure 5 plots the probability of successfully engaging five threats, or five attacks by a single threat, for several single-engagement probabilities of success. This analysis assumes that the single-engagement probability of success is fixed, that it is the same for all layers, and that the engagements do not interfere with each other. Though, in general, the probability of a successful engagement is not the same for all layers, Figure 5 illustrates the benefit of layering.

Figure 4. *Layered soft-kill.*



Coordination of EW with Hard-Kill Systems

To maximize their contribution, EW systems must operate in concert with the entire combat system, and our soft-kill EW assets must be coordinated with our hard-kill missile and gun system assets. The required EW integration with ownship's hard-kill systems can be addressed at three levels: interference control, data sharing, and engagements coordination. The primary functions associated with these levels of coordination are as follows:

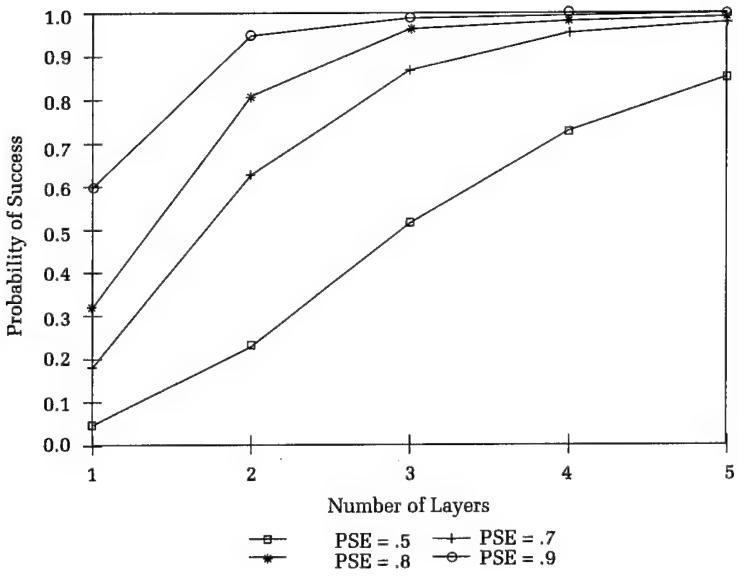


Figure 5. *Probability of successfully engaging five threats with multiple layers.*

Interference Control

- Real-time coordination of radar frequencies with ESM to ensure sufficient ESM look-through in threat bands.
- Real-time coordination of radar frequencies with ECM to ensure ECM does not interfere with radar performance.

Data Sharing

- Cuing fire-control and multi-function radars with ESM track data.
- Using ESM signal-amplitude data to support kill assessment for gun and missile systems.

- Using ESM emitter-type identification and platform correlation information in support of overall combat system threat identification.
- Using ESM tracking data to perform passive gridlock.
- Providing passive targeting for surface-to-surface missiles.
- Using ESM-emitter-mode data to assess tactical intent in support of hard-kill engagement decisions.
- Using radar data for the launching of decoys against targets not held by ESM.
- Using radar data in decoy placement algorithms and in support of time-sequencing of ECM techniques.
- Using radar track picture to help set ESM detection thresholds and other resource-management controls.
- Using radar track parameter and identification-friend-or-foe data to support emitter-type identification.
- Using radar kinematic track data to support ECM effectiveness assessment.
- Using radar tracking of decoys to support reseed logic.
- Using radar tracks to help resolve ESM de-interleaving ambiguities.
- Providing integrated tactical situational awareness.

Engagements Coordination

- Providing coordinated engagement doctrine.
- Using ECM to improve missile and gun system engagement geometry by forcing home on jam.
- Providing raid dispersion by counter-targeting and counter-locking on ECM.
- Selecting optimum engagement schedule to prevent saturation of both hard-kill and soft-kill weapons.

Correlation of ESM tracks with other combat system tracks is key to fully performing many of these functions. The angle-measurement accuracy of most current EW systems, though more than adequate for pointing ECM beams, is insufficient to support radar/ESM correlation in dense radar and ESM track environments. Optimally, an integrated EW system should have azimuth and elevation-angle accuracies similar to those of modern multifunction radars to keep reasonable levels of false correlation. As all tracks will not necessarily have an emitting radar, and because the detection range of radar and ESM can be quite different, the track files may contain substantially different sets of tracks. ESM/radar correlation will be limited to low to moderate densities without substantial improvements in current system track accuracy. The requirement to fully integrate hard-kill and soft-kill systems in our combat systems will be a major design driver in next-generation EW systems.

Future Tactical EW System Requirements

Future EW systems must be designed to operate in the RF environment of the next century, provide the needed layered EW defense, serve as a combat system sensor, and be fully integrated into our combat systems. Situational awareness will again become a major EW requirement, and next-generation EW systems will be capable of providing for long-range targeting for surface-to-surface missiles. The threat identification, assessment of intent, and situational awareness that a modern EW system can provide will play a major role in allowing our systems to operate in the quick-reaction modes required to counter the modern cruise missile threat. In many cases, ESM will be the sensor that initiates missile engagements. Greater emphasis will be placed on counter targeting, as these systems will begin the ECM battle well before the opposition launches its missiles. There will be greater reliance on active offboard decoys to counter the monopulse threat.

Certain performance parameters are key to achieving this increased combat system role. Table 2 lists these parameters and the requirements driving them.

There are a number of recent technology advancements that will enhance meeting these requirements. They include:

- Multiple element-array interferometers with high-angle measurement accuracy and modest antenna gain
- Channelized receivers
- Offboard active decoys
- Coherent repeaters using digital RF-memory technology
- Solid-state antennas and transmitters
- RF packaging improvements
- Monolithic microwave integrated circuits.

These technologies, when combined with the revolution in digital processing that makes the intense computational burden of processing and sorting millions of pulses per second possible, allow significant improvements to be made over today's tactical EW systems.

Table 2. *Performance Parameters and Requirements*

Performance Parameter	Driving Requirements
Sensitivity	• Long-range situational awareness • Over-the-horizon targeting • Side-lobe detection for counter-targeting and counter-missile lock-on • Tracking continuity through signal fades • Tracking of power-managed emitters • Detection of LPI threats
Frequency coverage	• Expected range of threat frequencies
Spatial coverage	• Expected threat profiles
Angle-of-arrival measurement accuracy	• Radar cuing • Over-the-horizon targeting • Radar/ESM track correlation • De-interleaving frequency and waveform agile emitters • Correlating multiple emitters to a single platform
Receiver channel bandwidth	• Performance in hostile ECM environments • Emitter density • Isolation from ECM • Electromagnetic interference (EMI) from other ownship emitters
Pulse throughput capability	• Emitter density
Number of simultaneous and overlapping pulses to be detected and measured	• Emitter density
ID Accuracy	• Over-the-horizon targeting • Situational awareness • ECM technique selection • Radar cuing
Number of simultaneous engagements	• Expected raid densities
Reaction time	• Counter targeting (especially side-lobe jamming) • Counter-missile lock-on (especially side-lobe jamming) • Late turn-on threats
ECM frequency range	• Threat
Effective radiated power	• Side-lobe jamming for counter-targeting and counter-missile lock-on • Missile burn-through range
ECM beamwidth and side-lobe levels	• Isolation between ESM and ECM • EMI to other combat system sensors

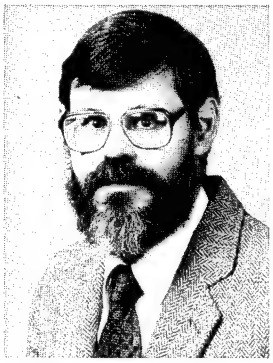
Summary

This article has briefly addressed the benefits EW can bring to ship defense and the direction EW must follow to do so. It can no longer be thought of as a stand-alone capability, but must be considered in the context of the entire combat system. Tactical EW is beginning a new era as the Navy begins to look at a broader contribution of our EW systems to our combat systems. Current efforts to define the next-generation EW system are addressing the performance requirements and threat trends identified in this article. The requirements definition for the FY 2003 combatant is addressing the full spectrum of EW contributions across all warfare mission task areas. Merging the management of our tactical EW systems with other ship defense systems in the Ship Self-Defense Office of the Theater Air Defense Program Executive Office is ensuring that EW is addressed as part of total ship self-defense.

References

1. Schleher, D. Curtis, *Introduction to Electronic Warfare*, Artech House, Norwood, MA, 1986.
2. Kiely, D. G., *Naval Electronic Warfare*, Brassey's Defence Publishers, London, 1988.
3. Schlesinger, Robert J., *Principles of Electronic Warfare*, Peninsula Publishing, Los Altos, CA, 1979.

The Author



THOMAS W. KIMBRELL is the Chief Engineer in the Electronic Warfare Division. He came to NSWCDD in 1969, after receiving a B.S. in electrical engineering from Tulane University. In his 25 years at NSWCDD, he has worked in the Warfare Analysis Department (now the Strategic Systems Department), in the Combat Systems Department, and the Ship Defense Systems Department in the areas of digital gun fire-control systems, real-time ballistics computations, combat system design, computer control of radar systems, phased-array radar systems, radar electronic counter-countermeasures design, and electronic warfare. Before becoming Chief Engineer, he was head of the Electronic Warfare Integrated Branch and, prior to that, he was head of the Radar Systems Branch. He is currently involved in defining requirements for the Advanced Integrated Electronic Warfare System and electronic warfare combat system integration.

Signal Simulators Used in Deception

T. Welle, C. Beatty, R. Kirby, T. Spradlin, C. Johnson, and A. Furano

Most signal synthesizers available today are expensive and limited in flexibility; that is, not software programmable. This article describes a Naval Surface Warfare Center Dahlgren Division (NSWCDD) program to develop a flexible and affordable high-fidelity signal generator using commercial off-the-shelf (COTS) subsystems. NSWCDD demonstrated that it is possible to simulate complex radar waveforms with a high degree of fidelity while eliminating many of the problems associated with older techniques. The synthesis of complex waveforms is essential for conducting research and developing modern weapon systems. These synthesizers will form the central element in modern tactical deception and countermeasure systems.

Introduction

Signal simulator technology is constantly affected by today's complex and ever-changing electromagnetic (EM) environment. Modern radars use frequency agility, staggered pulse repetition intervals (PRI), and compressed waveforms to increase the probability of detection and reduce vulnerability to jamming.

Of the signal simulator systems available today, most are expensive to develop and maintain, and are usually limited in flexibility. As signal requirements change, time-consuming and expensive modifications are required to accommodate the changes. Because of the time and expense involved in modifications, there is a broad interest in equipment that can be quickly and economically reconfigured to meet changing needs.

NSWCDD developed a radar demonstration unit (Figure 1) to simulate the EM signature of multiple complex radar emitters. Two pulse comparisons between a real and a radar demonstration unit simulated complex radar are shown in Figure 2. The upper trace is linear amplitude modulation (AM) demodulated video, and the lower trace is frequency modulation (FM) demodulated video. These signals were collected off-air from two different locations. A significant amount of waveform variation is due to environmental and geographical effects. A typical electromagnetic support measures (ESM) operator keys off the pulse width, pulse repetition interval, and the characteristics of the amplitude and frequency modulations. The radar demonstration unit software can be completely configured and consists primarily of COTS subsystems to facilitate installation and maintenance.

Approach

To realistically replicate an EM environment, a broadband simulator system must be capable of duplicating multiple emitters with complex waveforms. These complex waveforms include frequency-agile pulses that may be phase and frequency encoded. The simulator should also be capable of generating

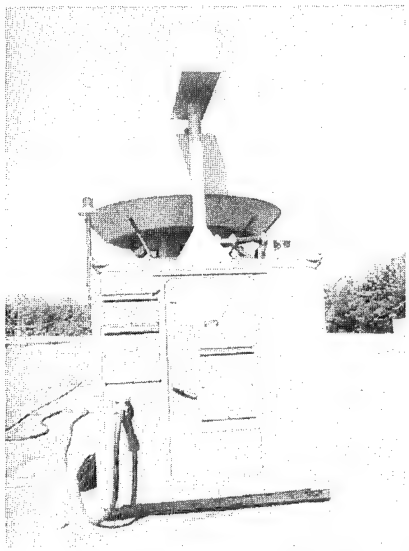


Figure 1. *Photograph of radar demonstration unit.*

high-fidelity antenna patterns, both in azimuth and elevation.

A graphical user interface was used to provide the ability to easily build and edit emitter, platform, and scenario files. Although these design requirements were difficult to meet, analysis conducted during preliminary planning indicated that recent advances in digital-synthesis-based signal simulators, specifically the Hewlett Packard (HP) Frequency Agile Signal Simulator 8791 Model 11, could provide the basis for the radar demonstration unit. The radar demonstration unit frequency-agile signal simulator is a COTS synthesizer capable of high-fidelity reproduction of pulse waveforms across a wide-frequency spectrum. However, it was not designed for multi-emitter operation, or to have the ability to replicate complex antenna scan patterns. An additional device, the digital environment generator, is required to implement these capabilities.

Radar Demonstration Unit

The four major subsystems of the radar demonstration unit include the frequency-agile signal simulator, digital environment generator, radio frequency (RF) distribution subsystem, and system controller (see Figure 3). The actual size of the radar demonstration unit is shown in Figure 4. The frequency-agile signal simulator generates the RF and intrapulse characteristics. The digital environment generator manages the emitters, and the interpulse and antenna scan characteristics for each emitter. The RF distribution distributes, amplifies, and radiates the low-level signal generated by the

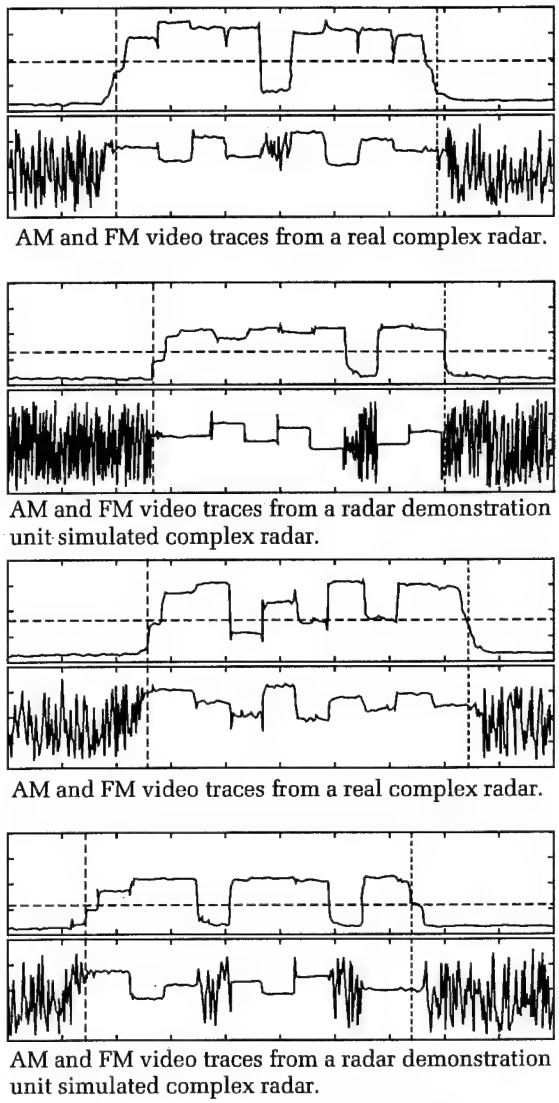


Figure 2. *Pulse comparisons.*

frequency-agile signal simulator. The radar demonstration unit can generate three independent complex pulse trains. Moreover, because the design is modular, new emitters can be easily added.

Frequency-Agile Signal Simulators

The frequency-agile signal simulator was developed as an economical alternative to traditional simulation methods. It can generate most of today's complex pulse waveform frequencies from 10 MHz to 18 GHz. The frequency-agile signal simulator is a unique combination of memory and sequencers coupled to a direct digital synthesizer driving a phase-coherent agile upconverter (AUC). The memory stores the sequences for controlling the carrier frequency (RF), FM, phase modulation (PM), and AM characteristics of the intrapulse waveform, enabling the frequency-agile signal simulator to generate

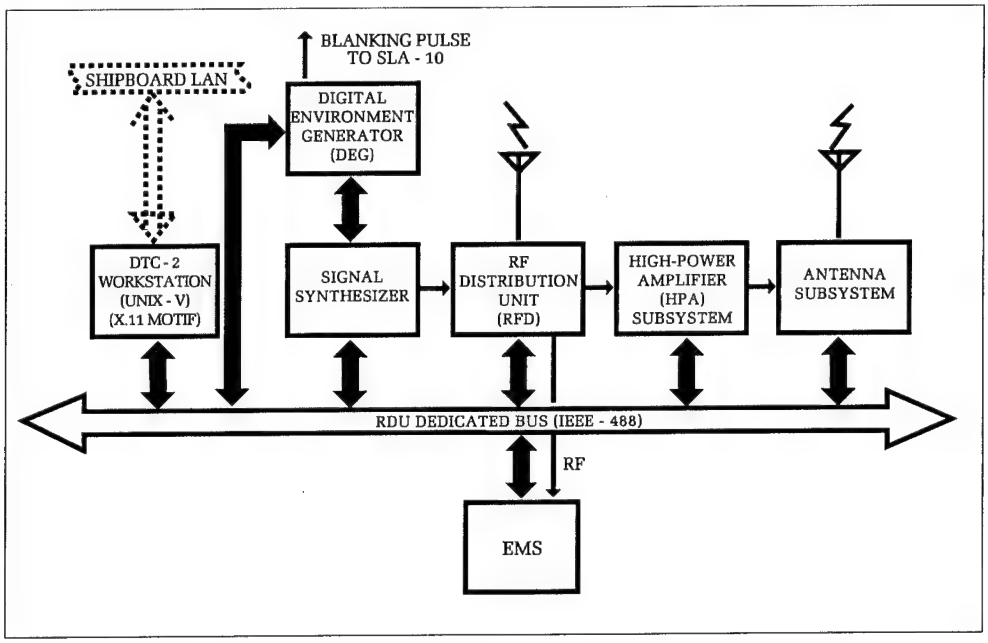


Figure 3. *Radar demonstration unit basic block diagram.*

frequency-agile, chirped, or phase-coded pulses.

The frequency-agile signal simulator comprises the modulation data source, the agile carrier synthesizer, the AUC, and the smart interface. The modulation data source contains the AM, FM, PM, and FREQ memories and sequencers used to produce the digital representation of the desired RF waveform. The agile carrier synthesizer converts the digital data from the modulation data source into an analog waveform in the 14 to 58 MHz frequency range and is also responsible for defining the instantaneous amplitude of the waveform. The AUC filters the 14 to 58 MHz agile carrier synthesizer output and upconverts it to 0.1 to 3 GHz. The AUC also contains amplifiers and attenuators to precisely control the signal output level. The smart interface is an 80286-based personal computer (PC) that combines the individual components into one unit within the frequency-agile signal simulator and allows the user to interface with the frequency-agile signal simulator while in stand-alone mode. The smart interface also routes data and

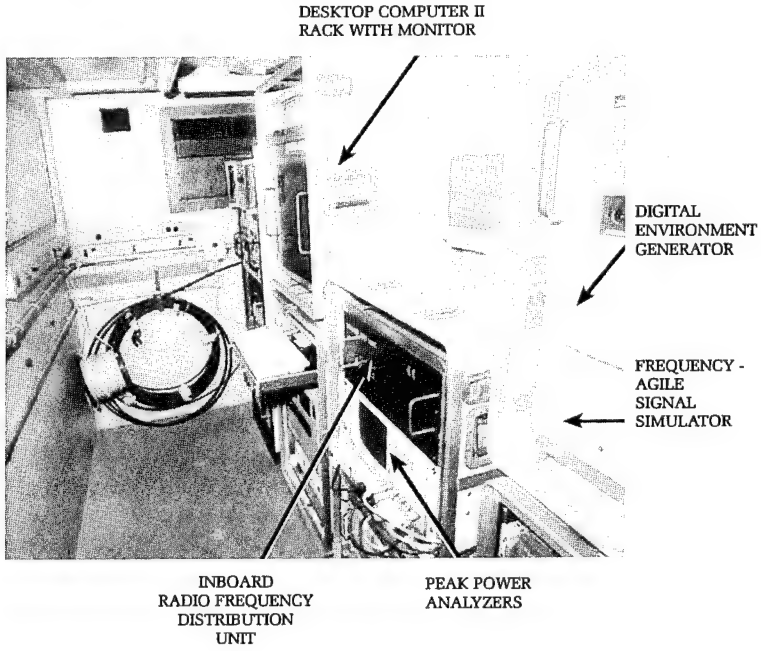


Figure 4. *Major components of radar demonstration unit inside van.*

commands to the correct component (modulation data source, agile carrier synthesizer, or AUC) within the frequency-agile signal simulator through the internal HP instrumentation bus.

Programming the Frequency-Agile Signal Simulator

Although simulating a radar is a complicated task, it can be simplified by breaking it into several smaller, manageable pieces. Mechanically steered antennas do not require simulation effort by the frequency-agile signal simulator, because only the antenna scan pattern (a digital environment generator responsibility) is seen by a far-field observer. However, an electronically steered antenna relies on a phase or frequency shift off of the main RF carrier to guide the radar beam. To simulate the radar of an electronically steered antenna, the frequency-agile signal simulator provides a PM or FM memory shift, and the shift is coupled with the digital environment generator generated antenna scan pattern. The point of view from which the simulation is observed in elevation is also important, because the phase (or frequency) shift required to accomplish a given elevation range will dictate how many of the radar's beams (or pulse types) need to be programmed into the frequency-agile signal simulator.

Because the digital environment generator is controlling the external attenuator to simulate the radar's antenna scan pattern, only one copy of each simulated radar pulse is needed in AM waveform memory. A sequence is programmed into AM sequencer memory and contains several packets. A packet is the actual "playback" of the values programmed into the waveform memory.

The sequencer/packet structure is required to be flexible enough such that, by using the values programmed into AM waveform memory, the pulse waveform can be described entirely in a few packets. For example, one packet can describe the rise, pulse on, and fall times of the waveform, or each packet can describe a component of the waveform (such as rise, pulse on, or fall time). Because there can be as many as 1024 frequency-agile signal simulator sequences, the digital environment generator selects a sequence and prompts the frequency-agile signal simulator to start execution. After a pulse has been defined in AM memory, the FREQ memories are programmed to generate the correct center frequency of the radar.

To simulate a radar that uses an interpulse (pulse-to-pulse) frequency change to scan electronically in elevation, the FREQ memory is programmed with the new RF frequency (refer to Figure 5). The radar scans in a bottom-top-bottom-top sequence, so the radar pulse shown in Figure 6 would be seen by an observer stationed on the horizon. Note that from the observer's position, only the first beam in the scan appears at maximum amplitude. Because beams 2, 3, and 4 are not pointed directly at the observer, only the side-lobe energy from these beams will ever reach the observer. Therefore, these beams will appear to be at a lower amplitude relative to the first beam, which is aimed directly at the observer.

To simulate this effect, the frequency-agile signal simulator AM memory is programmed with amplitude modulation on the pulse (AMOP). The AMOP levels for beam power at the observer's position are calculated from the antenna beam widths of the real radar. The observed AMOP from an actual radar is caused by a frequency change to electronically steer the beam. A simulation of the radar must provide AMOP as well as the corresponding frequency shift. The FM memory is programmed with the correct offsets so that the combination of the AM and FM sequences produces a frequency offset under the pulse.

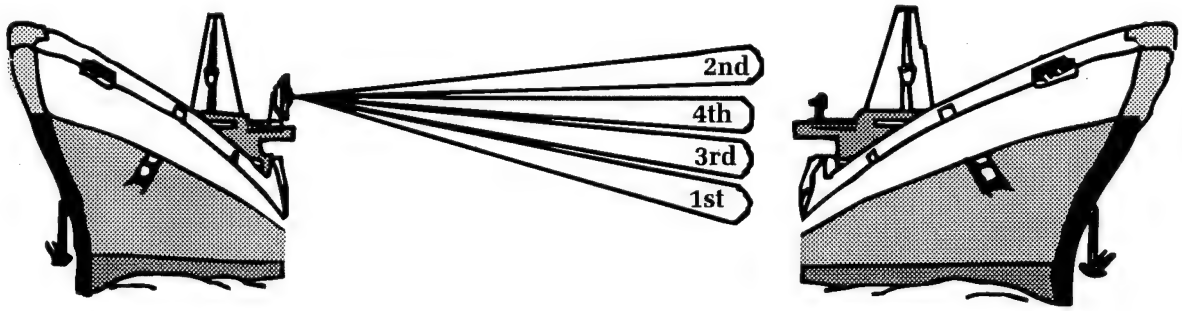


Figure 5. *Radar scan sequence.*

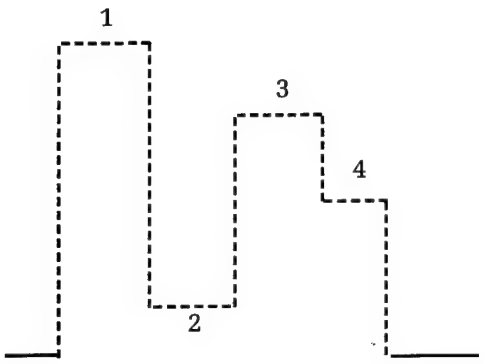


Figure 6. *Observed radar pulse.*

Digital Environment Generator

The digital environment generator is capable of generating pulse trains with up to 512 levels of stagger and pulse jitters up to the PRI. The digital environment generator can simulate complex antenna scan patterns using captured off-the-air radar signatures. Using this technique, pattern permutations unique to a specific ship can be simulated. The digital environment generator can also replicate the unintended scan modulations generated by the roll and pitch of a ship.

The digital environment generator is designed around a durable 80386-based PC chassis (refer to Figure 7) and provides all required functions except the PRI and SCAN processing, which are provided by digital signal processing (DSP) cards. The digital environment generator communicates with the system controller and other subsystems through an internal IEEE-488 interface board. An EEPROM board provides nonvolatile memory.

The digital environment generator interfaces with the frequency-agile signal simulator, RF distribution, system controller, and the shipboard blanking system (SLA-10). During simulation, the digital environment generator sends a digital word to the frequency-agile signal simulator to select the desired pulse sequence, followed by a trigger at the appropriate time. The digital environment generator interfaces with the RF distribution to provide discrete control for the RF switches, attenuators, grid pulse distribution to the high-power amplifiers, and a power meter trigger. A pulse is provided to the SLA-10 to blank shipboard receivers during the simulator's high-power transmissions.

A single DSP PC card performs the scan processing for all emitters in the simulation. A separate DSP PC card is provided for each unique emitter simulated. The DSP cards (COTS DX Products, DXP-25) use a Texas Instruments TMS320C25 fixed-point digital signal processor CPU. This board has 64 Kbytes of onboard RAM and operates with a 36-MHz system clock. This DSP card was chosen for its fast arithmetic calculation capabilities. NSWCDD designed and

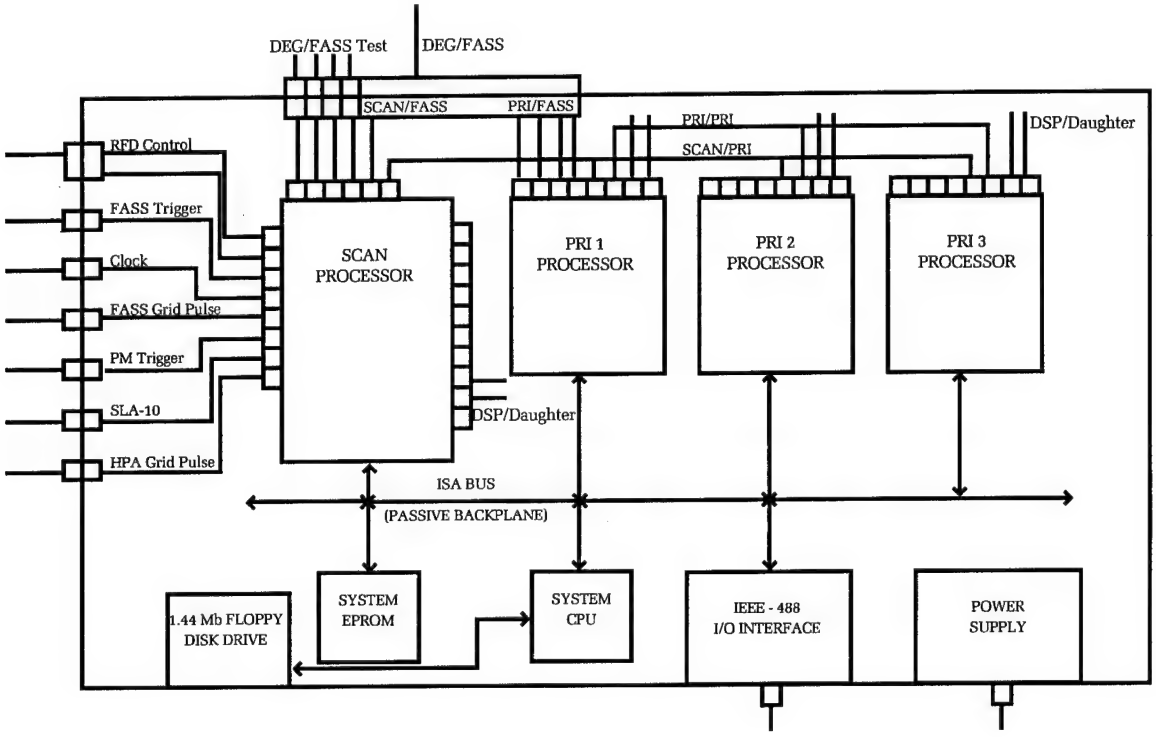


Figure 7. *DEG board level block diagram.*

produced prototypes of the daughter boards that provide the unique internal and external interfaces.

The PRI board, through the frequency-agile signal simulator interface, selects a pulse for a particular emitter and the proper frequency for simulating the ship's pitch and roll. An emitter's pulse interval is loaded into the internal timer on the PRI board. While the timer is counting down, the next pulse interval and associated jitter are calculated. The SCAN board determines which PRI board is enabled and signals the frequency-agile signal simulator to latch the data at the AM, PM, FM, and FREQ ports.

Because the frequency-agile signal simulator can generate only one RF pulse at a time, only one PRI board at a time may request frequency-agile signal simulator services. The SCAN board arbitrates PRI requests when pulse coincidence occurs. To signal the frequency-agile signal simulator to generate a pulse, the PRI board sends a pulse to trigger an external interrupt on the SCAN board through the SCAN daughter board. The SCAN daughter board has the capability of blocking the PRI-board triggers when a higher priority pulse is being generated. The SCAN board selects which PRI board output latches are enabled and triggers the frequency-agile signal simulator to generate the selected RF pulse. The values written to the frequency-agile signal simulator's AM, PM, FM, and FREQ interface ports are then calculated and loaded into the PRI-board output latches between each pulse interval.

The arbitration method used in the SCAN board for selecting a PRI board and the associated attenuation values is divided into two steps—amplitude arbitration and user-selected emitter priority arbitration. Amplitude arbitration is accomplished by separating the scan pattern into four amplitude regions. Regions with the highest amplitude have higher priority, and emitter(s) with amplitude values falling in this region are selected. If more than one emitter has identical amplitude values, the emitter with the highest user-assigned priority is selected for generation.

An arbitration look-up table is used to complete the two-step arbitration in one clock cycle. Pulses for emitters(s) that lose out in the arbitration are dropped. Pulse dropout studies of the various side-lobe levels and the two-step arbitration process were performed. The results of a 16-emitter simulation showed that an average of 1.2 percent of the pulses was dropped. Because of the arbitration method, the pulses with the highest amplitude (main beam) are almost never dropped.

Several sources of incidental (unwanted) jitter were discovered in the digital environment generator, and two of these sources were subsequently eliminated. The remaining sources of incidental jitter are the result of the digital environment generator's asynchronous design. To eliminate incidental pulse-interval jitter by the PRI boards, the PRI-board CPU idles after completing the calculations and waits for an internal timer interrupt. To eliminate the incidental jitter caused by the processing of an external interrupt on the SCAN board, the PRI board trigger also resets a SCAN daughter-board counter. This counter is read when the SCAN board CPU starts processing the external interrupt, which reduces all but one clock cycle of incidental jitter. The maximum amount of time it takes to start processing the interrupt is subtracted from the actual time taken to start processing. The remaining time is consumed by executing a series of no-operation instructions to keep the interrupt processing time constant. Before the PRI-board CPU idles, the next timer countdown value is calculated.

When the PRI board timer interrupt occurs, the precalculated timer value is read by the timer, and the countdown is restarted. The determination of these values in software allows for quick algorithm implementation and adjustment. An emitter's PRI, including pulse-interval jitter, is calculated using the following equation:

> Timer Countdown Value = Modified PRI Value = Loaded PRI + Remainder [(Random Number) / (Maximum Allowed Emitter Jitter Level)]

The current implementation employs a random number table with 32,000 pregenerated random numbers. The random numbers are divided by the maximum allowable jitter level, and the remainder is used as the PRI jitter offset value. This modified PRI value becomes the next timer countdown value. A follow-on to the prototype digital environment generator could employ a real-time random number generator, which would allow for more variations and flexibility in generating the PRI jitter.

The SCAN board selects a scan attenuation value, and the corresponding PRI board output is latched and triggers the frequency-agile signal simulator to read the interface ports. The SCAN board also sends the scan attenuation value to the RF distribution. Other functions of the SCAN board include the timing and distribution of the grid pulse and RF distribution RF switch control. Digital PIN attenuators in the RF distribution, which are controlled by the SCAN

board, generate the scan pattern for each emitter.

The system controller provides the SCAN board with 4096 attenuation values representing an emitter's 360-degree scan pattern. The internal SCAN board CPU timer is set to time out every 250 microseconds and is used to increment through the attenuation values. This implementation results in scan rate selections of 1.024 seconds, 2.048 seconds, 3.072 seconds, etc. The attenuation amount for each emitter, including scan pattern noise, is calculated between timer interrupts.

To prevent the scan pattern from looking too "clean" to an ESM operator, random scan noise is added to the generated scan pattern. Because the noise level varies with the scan amplitude, an attenuation look-up table with 1024 values covering 0 dB to 64 dB of attenuation is used to determine the maximum amount of noise allowed at a specific point in the scan pattern. The level of noise increases as the attenuation values increase. A scan noise look-up table containing 10,000 randomly generated numbers is used to determine whether the scan noise is added to or subtracted from the attenuation values. If the least significant bit of the random number is 0, scan noise is added. If the least significant bit of the random number is 1, the scan noise is subtracted. The random numbers are divided by the maximum allowed noise level, and the remainder is used as the noise offset value.

The attenuation value for each emitter is used to control the setting of the digital attenuators and generation of the arbitration table address. The attenuation value is calculated using the following equation:

> Scan Attenuation Value = Loaded Attenuation + Remainder [(Random Number) / (Maximum Allowed Scan Noise Value)]

The digital environment generator firmware currently allows for the termination of all pulses. Before the frequency-agile signal simulator has completed the generation of a pulse, another request may be made by another emitter (PRI board trigger). If the SCAN board triggers the frequency-agile signal simulator while the frequency-agile signal simulator is processing a pulse, the frequency-agile signal simulator terminates that pulse and begins processing the new pulse. The SCAN daughter board has the capability to block the termination of a pulse. The SCAN board requires the pulse length and termination status data for each pulse from the system controller to implement this capability.

Tradeoffs are made when generating asynchronous or synchronous multiple-emitter, pulse repetition intervals. These tradeoffs concern the generation of incidental jitter and synchronized pulses from more than one emitter. Because the PRI boards, the SCAN board, and the frequency-agile signal simulator have separate clocks, a total of 253 nanoseconds of incidental jitter is introduced by the digital environment generator. The PRI and SCAN board interfaces and the digital environment generator frequency-agile signal simulator interface are responsible for the creation of this incidental jitter. The incidental jitter can be eliminated if the digital environment generator uses a master clock supplied by the frequency-agile signal simulator. The disadvantage of this configuration is that all pulse trains generated by the system will be synchronized, which may be detected by some ESM operators.

A compromise may include synchronizing the frequency-agile signal simulator, the SCAN board, and some of the PRI board clocks. The remaining PRI boards would have the capability of generating asynchronous emitters. The emitters selected for generation by the asynchronous PRI boards would normally generate a minimum of 253 nanoseconds of incidental jitter.

Radio Frequency Distribution

The primary purpose of the RF distribution is to distribute, amplify, and radiate the low-level signal generated by the frequency-agile signal simulator. RF distribution components are located in an internal shipboard space and in topside enclosures. The RF distribution was designed to allow for computer-controlled diagnostics and calibration functions. The RF distribution contains a peak-power meter that allows the system controller to sample the RF power at any point along the RF path.

One of the advantages of this design is that an end-to-end calibration can be performed to maximize the dynamic range and to remove all nonlinearities in the system. Nonlinearities are caused by the distribution amplifiers, filters, and the high-power amplifiers, in the RF chain.

Radio Frequency Subsystems

The RF subsystem comprises an inboard RF distribution unit, two outboard RF distribution units, a peak power meter, a traveling wave tube amplifier, two half-omni antennas, and numerous RF connections (see Figure 8). To radiate the signals generated by the

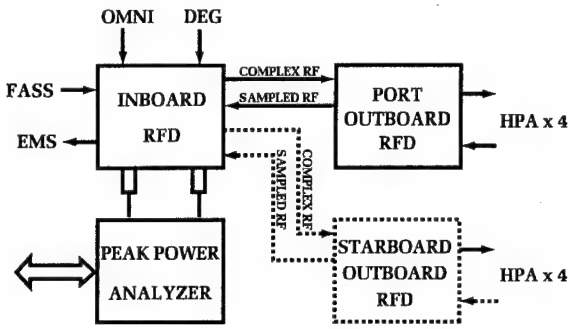


Figure 8. *Radar demonstration unit RF subsystem block diagram.*

frequency-agile signal simulator and digital environment generator, the RF subsystem is required to

(1) amplitude modulate the signal for antenna scan pattern generation,
(2) conduct the RF signals from below deck space to both port and starboard antennas/amplifiers unit above deck,
(3) calibrate for RF subsystem nonlinearities,
(4) amplify the signal for radiation, and
(5) radiate the signal 360 degrees.

The radar demonstration unit simulates an RF pulse and an antenna pattern's main, first, and second side lobes. Because the RF signal generated by the frequency-agile signal simulator is at a constant amplitude, digital environment generator controlled, PIN diode attenuators in the inboard RF distribution unit amplitude modulate the signal. For example, if an emitter has a second side-lobe level 26 dB below the main beam, the digital environment generator will issue a command to the PIN diode attenuators to add 26 dB of attenuation to the signal. Interpulse AM is also done in this way. The PIN diode switching noise is eliminated, because switching is done when there is no signal present, and the traveling wave tube amplifiers are not transmitting.

The majority of the radar demonstration unit equipment resides in a controlled environment below deck. A port and starboard outboard unit are placed topside and can be placed up to 200 feet away from the equipment below deck. Each outboard unit is composed of a traveling wave tube amplifier, an outboard RF distribution unit, and a half-omni antenna.

The inboard RF distribution unit located below deck compensates for transmission cable losses to the outboard units by amplifying the signal with a single wide-band (0.1-18 GHz) amplifier. The amplified signal is then split and transmitted to the port and starboard outboard units through low-loss, wide-frequency bandwidth cables. Each outboard RF distribution unit amplifies the signal again and, using a special quadriplexor, splits the RF signal into four frequency bands before sending it to the traveling wave tube amplifiers. The quadriplexor and up to four traveling wave tube amplifiers were used, because a single traveling wave tube amplifier with the required power, frequency bandwidth, pulse width, and duty cycle was not available.

Pulsed traveling wave tube amplifiers are used to amplify the RF, because other COTS high-power amplifiers did not have the required frequency bandwidth. For the radar demonstration unit, only one traveling wave tube amplifier was used because of cost restraints. Duty cycle, pulse repetition frequency (PRF), pulse width (PW) range, and pulse stability specifications were generated to support the duty cycle, PRF, and PW of the emitters in the frequency band. Also, traveling wave tube amplifier noise power output with no input signal present had to be very low. Other concerns were the size and weight of the traveling wave tube amplifiers.

Because the radar demonstration unit is required to radiate an RF signal 360 degrees, port and starboard antennas are used, each having an antenna pattern of 180 by 10 degrees with a slant polarization of 45 degrees. Because most receiver antennas are either horizontal, vertical, or circular polarized, the 45-degree slant polarization allows the radar demonstration unit's RF signal to be collected by any receiver with minimal signal loss. For example, the use of a vertically polarized transmitter antenna and a horizontally polarized receiver antenna would result in the received signal being attenuated over 20 dB. The use of a 45-degree slant, polarized transmitter antenna and the same receiver antenna would result in the received signal being attenuated only 3 dB. An antenna meeting these requirements was not found; as such, the development of an appropriate antenna was performed at NSWCDD. A horn antenna with a special feed and overextended flares was developed.[1]

RF System Calibration

To maximize usable dynamic range, RF system calibration compensates for any nonlinear characteristics inherent in the RF signal path and allows the nonlinear portion of the traveling wave tube amplifiers' output near saturation to be used. The system

controller calibrates the RF system before the system is operational using calibration software that controls the frequency-agile signal simulator, peak power meter, inboard RF distribution units, and outboard radar demonstration units. Calibration results are stored by the system controller and are used to linearize, relative to the RF subsystem, the emitter scan pattern data that is downloaded to the digital environment generator.

The frequency-agile signal simulator generates a special calibration pulse that is passed through the entire RF chain and is measured by the peak power meter after the traveling wave tube amplifier. Initially, the personal identification number (PIN) attenuators are set for no attenuation of the calibration pulse, thus driving the traveling wave tube amplifier into saturation. The resulting power output is measured, and the attenuation provided by the PIN attenuators is increased by 1 dB. This process is repeated until a peak power measurement is taken that is slightly less than the previous one. At this point, the calibration program determines that the peak power threshold has been passed and adjusts the PIN attenuator setting to the midpoint of the last two power measurements. This is considered to be the peak power setting for this particular traveling wave tube amplifier.

After finding the peak power setting, the traveling wave tube amplifier nonlinear characteristics near saturation are determined by using the calibration pulse and increasing the attenuation of the PIN attenuators in 1/4-dB increments. After each increment, the peak power meter measures the traveling wave tube amplifier output and reports the value to the desktop computer (DTC). This process is repeated until the traveling wave tube amplifier output is down 15 dB from the peak power and is thus in the traveling wave tube amplifier's linear output region.

The system controller uses this process and performs a linear interpolation to construct a 1024-element sequential table of PIN attenuator settings required for the desired output levels. The first entry in the table is the PIN attenuator peak power setting, and each entry after that is the attenuator setting required to produce an end power level change of -1/16 dBm. The system controller uses this calibration table when downloading the digital environment generator. Each value in a digitized emitter scan pattern, 0 being maximum power and 1023 being no output, is used as the index into the calibration table. The PIN attenuator setting located at the index is the value required for the digital environment generator to generate that particular part of the scan pattern.

System Controller and User Interface

The system controller is a second-generation U.S. Navy standard DTC-2. The DTC-2 is a COTS UNIX durable workstation based on the SUN scalar processor architecture (SPARC). The system controller contains both a general purpose instrumentation bus (GPIB/IEEE STD-488) for communication with all other radar demonstration unit hardware, and an ethernet interface for communication with shipboard systems connected to the shipboard local area network (LAN).

Because the radar demonstration unit is scenario driven, the operator must use the system controller to create scenario descriptions that drive the simulation. The system controller manages the emitter, platform, and scenario data bases; and executes scenarios using a mix of both COTS (e.g., data-base management system) and custom software. Custom software applications were developed using X-windows and the OSF/MOTIF Graphical User Interface (GUI) to provide system specific functionality such as the system monitor (basic system control function), scenario execution, and system built-in test (BIT).

The software system comprises a set of applications, each designed to provide a specific software function or hardware interface. Each application handles a well-defined set of messages. Data exchange between applications is handled with interprocess communication. Figure 9 illustrates the high-level software architecture and the interface to hardware.

The application software was developed using a standard application framework that provides a common kernel functionality as well as a common message interchange format and handling mechanism. These characteristics allow for application portability, reduced development time, and a reduction of software maintenance costs.

The message-passing software design distributes applications between various systems on a network. For example, the user interface application can be located on a work station in the Combat Information Center, while the hardware control functions are located on the system controller in a radar space. This allows for a very flexible and reconfigurable software-control environment.

During the radar demonstration unit effort, a standard application framework was developed that provides a common foundation on which all applications in the software system are designed and implemented. The standard application framework encompasses a standard for application software structure,

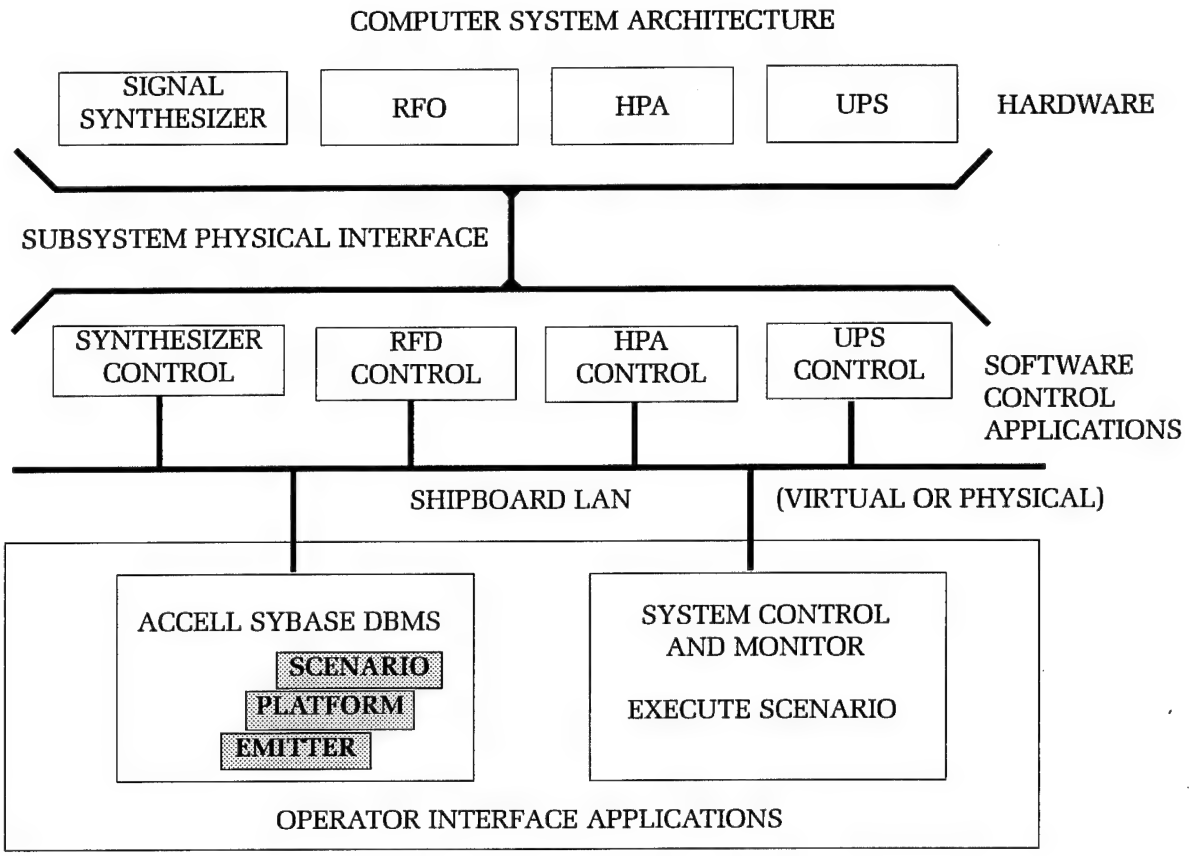


Figure 9. *System controller architecture.*

standard system-level functionality, and a standard communications protocol that all applications use to interact. The standard application framework is documented in detail in a white paper.[2]

External interfaces to the system controller include the GPIB/IEEE-488 and RS-232. Internal interfaces rely on the standard application framework standard message-handling protocol to provide the necessary interprocess communication. Each of the internal interfaces are defined by the messages that each service handles.

System Controller Functionality

The system controller software consists of several major functions including:

- Emitter Library
- Scenario Library
- Platform Library
- Execute Scenario
- BIT
- System Monitor

Emitter Library

The emitter library function permits the operator to view the available radar demonstration unit emitter description files and the emitter parametrics used to create those radar demonstration unit emitter description files. The emitter library data base contains emitter parametric data for each platform-specific emitter/mode simulated by the radar demonstration unit. The emitter parametric data comprises an emitter descriptor that is the identifier of the emitter to be radiated between specified start and stop times. The emitter description file contains the frequency-agile signal simulator and digital environment generator binary load data as well as the SCAN and PRI data for the digital environment generator.

Scenario Library

The scenario library function allows the operator to develop radar demonstration unit scenarios that describe the operation of simulated radar emitters. Each scenario is identified by a scenario descriptor and contains an emitter descriptor for each emitter in the scenario, an emitter priority for each emitter (relative to other emitters in the scenario), a scenario duration time required to perform the simulation (hours:minutes:seconds format), and emitter events (stop and start times of each emitter in the scenario). All scenario times are specified relative to the start of

scenario execution (relative to time 00:00:00), are recorded in hours:minutes:seconds format, and are accurate to within 1 second. The scenario library also allows the user to set up operator events that notify the user of a required action. Operator events are identified by an operator event descriptor and contain a start time relative to time 00:00:00 and a cue message to be displayed to the operator describing a required action on the part of the operator.

All the information specified above is compiled into a scenario execution file that contains the load emitter, start emitter, and stop emitter messages used to control the digital environment generator and the frequency-agile signal simulator during scenario execution.

Platform Library

The platform library function allows the user to maintain a data base of platforms that may be used during the execution of scenarios. This data base describes each platform's name, the platform type, emitters commonly found on that platform, and any physical or dynamic (pitch/roll) characteristics.

Execute Scenario

The execute scenario function executes the scenario as described by the scenario description file and allows the user to monitor the progress of the scenario or to control the scenario execution. During scenario execution, the execute scenario function generates messages to the applications that affect radar demonstration unit hardware control in order to correctly perform the simulation and provide the operator with scenario status information, such as active emitters and scenario sea state.

At scenario execution time, the user supplies the current sea state. This value is used to determine the pitch and roll characteristics for the scenario simulation.

Built-In Test

The system controller provides a BIT capability for the entire radar demonstration unit system. All devices that are capable of providing status information and/or built-in-test functionality are exploited by the BIT function.

System Monitor

The system monitor module is the central point of control and monitoring for the system controller. It provides functions for system startup, operational readiness check, BIT, system status, and manual system control. This module allows the operator to monitor the system status or to manually control various devices in the system remotely via the system controller.

Conclusion

The objective of this project, to develop a flexible and affordable high-fidelity signal simulator, was accomplished. Our work suggests that it is possible to simulate multiple, complex radar waveforms with a high degree of fidelity, while eliminating many problems associated with older simulation techniques. The signal simulator used in the radar demonstration unit is affordable and very flexible, requiring only the development of a new parameter file in order to synthesize a new waveform. This achievement has prepared the base architecture and technology for modeling and simulation applications in the areas of radar, electronic countermeasures, electronic counter-countermeasures, guidance, and communication system analysis and design.

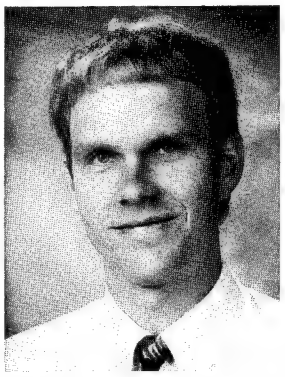
References

1. *I- and E/F-Band 180-Degree Azimuth (Elevation Confined) Antenna Research, Analysis, Brassboard Design and Evaluation*, Naval Surface Warfare Center Dahlgren Division, Revision 01, Aug 1992, CID 982047.
2. *The Standard Application Framework, Programmers Reference*, Naval Surface Warfare Center Dahlgren Division, Revision 1.3, Aug 1992, CID 982031.

The Authors

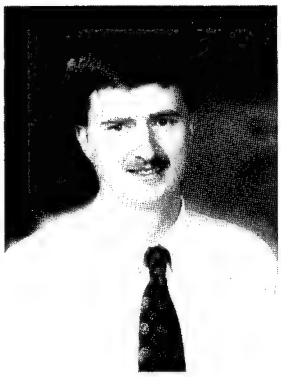
THOMAS N. WELLE received his B.S. degree in Electrical Engineering from Saint Cloud State University in 1987. He has been at NSWCDD since 1988, where he has worked in Electronic Warfare on the EP-3E ARIES II, CV-4019 Scan Converter, SCADS Radar Demonstration Unit, and the Cooperative Engagement Capability programs. He is a member of the IEEE.

CHARLES "CHICK" BEATTY received his B.S. degree in Electrical Engineering from the West Virginia Institute of Technology in 1970. He has been at NSWCDD since 1972, where he has worked in the areas of electronic warfare and intelligence. Mr. Beatty has been associated with the EMPASS, SLQ-32, AEDAS, DEEPWELL, ARIES, SCADS, and CEC programs. He is a member of the IEEE and the SAE.

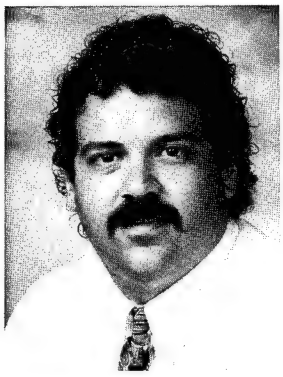


ROBERT H. KIRBY received a B.S. degree in Mathematics from Virginia Commonwealth University in 1990, specializing in computer science and operations research. He began working for the Naval Surface Warfare Center's Information Systems Branch as a co-op in 1987 and was converted to a Computer Scientist position upon graduation from VCU. Mr. Kirby has worked on a variety of projects including the SCADS radar demonstration unit, Tomahawk data analysis tools, and the Aries II airborne ELINT project. He is currently working on test of the Cooperative Engagement Capability's interfaces with various radar systems.

TOMMY SPRADLIN joined NSWCDD's cooperative education program in 1976 and received a B.S. in Electrical Engineering from Virginia Polytechnic Institute and State University in 1981. Since completing the B.S. degree, he has worked on various projects encompassing telecommunications, automated signal search and acquisition systems, EP-3E ARIES II upgrade, CIWS upgrade, radar simulation, sensor systems integration, and data fusion. Most recently his work has been in technology assessment for the 21st Century War Fighting.



CHARLES D. JOHNSON received a B.S. degree in Engineering Science and Mechanics in 1985, and an M.S. degree in Systems Engineering in 1989, both from Virginia Polytechnic Institute and State University. Since joining NSWCDD in 1985, he has performed as an Electronics Engineer on a variety of projects involving TCP/IP networking, Unix system programming, real-time multiprocessor system development, and X-windows application development. Specific project experience includes the ARIES-II Airborne EW system, SCADS Radar Demonstration Unit, Multisensor Detection, and the Cooperative Engagement Capability.



ANTHONY "TONY" FURANO received his B.S. degree in Electrical Engineering from The University of Cincinnati in 1985. As an undergraduate at UC, he co-opped at Cincinnati Microwave, where he designed circuits for RF detection and processing. In 1985, Tony began working at NSWCDD. While at NSWCDD, he has worked in Electronic Warfare on SLQ-32, EP-3E ARIES II, CV-4019 Scan Converter and SCADS programs. He is presently working on developing the next generation computer for the Shipboard Gridlock System.

BOGHAMMER—Make My Day! Phalanx Enters the Surface Warfare Arena

LCDR S. A. Borland

The analysis following the Gulf War of 1991 revealed an Achilles' heel in the might of the U.S. Navy. Large surface warships could be disabled by the primitive, yet effective, threat from a fast patrol boat armed with explosives or by fanatics with hand-held weapons. The answer lay with the Phalanx Close-In Weapon System (CIWS), which could be adapted using nondevelopmental items (NDI) of equipment to meet this threat and the threat posed from low-performance aircraft. However, to incorporate off-the-shelf equipment into a mature system and adapt to a new role requires a thorough knowledge of the capabilities of each NDI component and its performance in the hostile environment adjacent to a rapid-firing gun system. Experimental work at the Naval Surface Warfare Center Dahlgren Division (NSWCDD) examined the performance of various electro-optic (EO) devices and the lethality of gun systems against new targets. This work is pivotal to the concept of using NDI and is crucial in obtaining the best value for the Navy's money.

Introduction

The loss of the Israeli ship ELIAT to a Soviet-made antisurface missile during the war of 1967 proved the vulnerability of warships to this emerging breed of weapon. The realization that a major surface asset could be disabled by a relatively inexpensive and easily delivered missile made the U.S.—with the largest fleet known to the world—sit up and take serious notice. The answer lay with Phalanx (Figure 1), a rapid-firing Gatling gun pointed by a sophisticated radar system and controlled by computer, which provides a final layer of hard-kill defense. The Phalanx proved an immense success, with over 750 systems built for the U.S. Navy and several allied navies.

In the same way that the events of 1967 paved the way for the introduction of Phalanx, an analysis of the Gulf War of 1991 also highlighted the need for a weapon system to protect ships from a new threat. Rather than a high-speed delivery from the air, the new threat could be delivered at a slower speed—in the form of a small, manned boat packed with explosives or hand-held weapons. The Swedish-built BOGHAMMER fast boat was a prime example. This boat had been procured and put to devastating use by the Iranians during the war with the Iraqis. Just as in 1967, powerfully armed warships could be knocked out by an unsophisticated, yet efficient, threat simply because their weapons were not tuned against that one particular adversary. Again, the U.S. Navy realized this shortcoming and initiated studies to investigate a means for

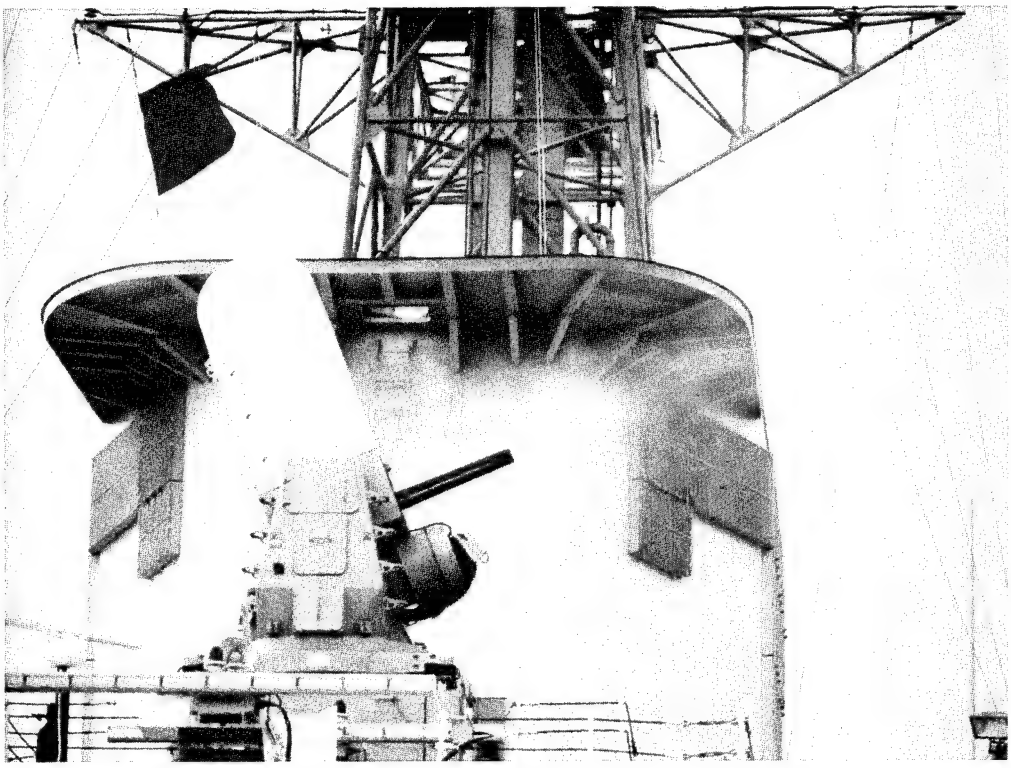


Figure 1. *The Phalanx Close-In Weapon System.*

countering this new threat. Accordingly, the Advanced Minor-Caliber Gun System (AMCGS) program was initiated to acquire a weapon system to combat the surface threat posed by a fast-moving small boat. At the same time, the Stabilized Weapons Platform System (SWPS) program was examining a system to combat the low-performance, close-range air threat. Late in 1991, a third program was initiated to determine whether a single weapon system could satisfy both AMCGS and SWPS requirements.

Why Phalanx Was Chosen

The program examined systems with sufficient firepower to rapidly defeat a threat that may reveal itself only at the last close-range moment. The new weapon should be available at all hours and in all weather. The U.S. Navy had such a system already standing guard over the fleet, one that may not require significant further development nor additional manpower to implement (major concerns in these days of financial constraint). The AMCGS/SWPS Cost and Operational Effectiveness Analysis revealed that, with only minor modification, Phalanx could counter this new surface threat. Moreover, it could simultaneously maintain its traditional role as defense against the antiship air threat. The importance of Phalanx as part of self-defense has also ensured its incorporation into the Ship Self-Defense System, currently being developed to combine elements of a ship's weaponry into a coordinated system working in harmony, rather than isolation.

The Surface-Mode Mission

The new mission for Phalanx was, by no means, simple to achieve. Not only must the system defend the host ship against small, high-speed, very maneuverable surface threats; it must also counter slow-speed air targets (such as the Cessna or a helicopter), which were deliberately designed to be beyond the reach of Phalanx. Despite this new capability, the primary mission of defending against the (now more substantially advanced) antiship missile air threat must be retained and must always predominate, even during surface engagement. This is a stiff order in itself, but to ensure cost effectiveness in these days of value for money, maximum use was also directed.

The operational environment in which the host ship could find itself was also far divorced from the original concept of *blue water* operations. Littoral warfare, close to land and with both friendly and hostile surface craft evident, is already a real and more

common operating environment, as the recent operations involving Somalia and Bosnia demonstrate. These criteria all differ from the original concept of Phalanx and, with the challenge laid down, the Division, acting as the Technical Direction Agent (TDA) in partnership with the Hughes Missile Systems Company (HMSC) as Design Agent, set to work.

The Original Phalanx

Developed by General Dynamics Pomona Division (now part of HMSC) in the 1970s, Phalanx is a self-contained defense system requiring only power and cooling water from the host ship. Apart from the controlling panels, there is no below-deck installation, making it relatively easy to install in most ships, given sufficient deck space. Totally autonomous in operation, all Phalanx components are self-contained (Figure 2). The radar servo structure (RSS), a search radar antenna mounted beneath a radome atop the upper system, maintains a lookout for potential threats. Painted white to deflect heat away from internal electronics, some systems have even been dressed to resemble snowmen during the wintertime; but snowmen do not pack a six-shooter as part of the normal inventory! The RSS moves in azimuth and elevation to bring the tracking radar to bear upon the target identified by the search radar. As the gun is also mounted on the RSS axis, the point-of-aim is near the line-of-sight of the tracker, thus alleviating the normal difficulties of alignment associated with displaced trackers and weapon systems. The gun is the proven M61A1 six-barreled Gatling, able to deliver 4500 rounds per minute of saboted tungsten projectiles.

Once a target has been detected, the onboard computer predicts the future position of the target and, given known criteria, determines whether the target poses a threat. Essentially, the incoming target must pass within a certain box around the ship and exceed a certain speed in order to be detected. It is for this reason that small boats and slow-moving aircraft can approach the ship without being engaged. Should the target meet the engagement criteria, firing would commence to achieve hits at the most effective range. This is extremely important, as many antiship missiles are supersonic and require a maximum number of hits over a short engagement window. Without this capability, even with destruction of the target, high-speed debris could still litter the ship, causing enormous damage.

To maximize engagement efficiency, Phalanx spots its own outgoing projectiles, determines the error between target and line-of-fire, then corrects the point-of-aim to reduce the error to zero. This concept, known as continuous aim correction, is vastly superior to the conventional prediction ballistics associated with other gun systems and lies at the heart of Phalanx's firepower efficiency.

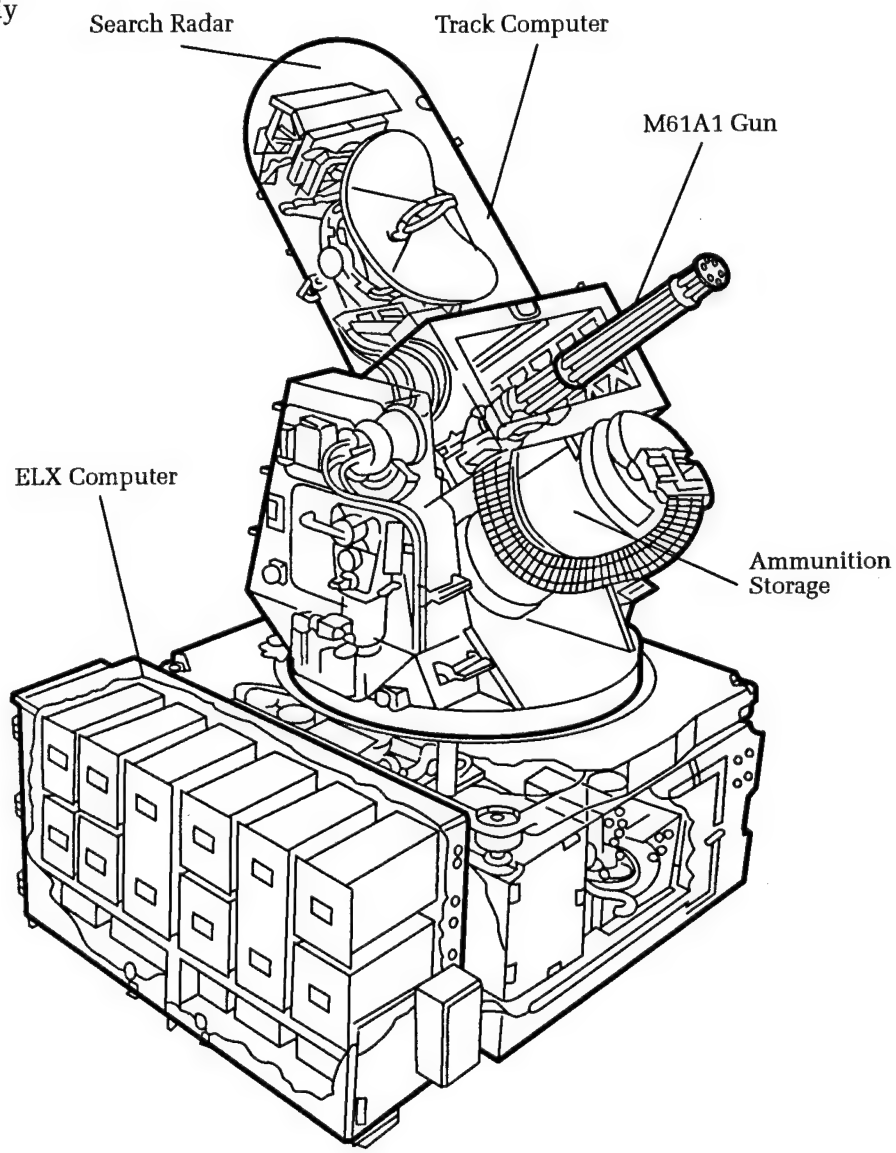


Figure 2. *Phalanx cutaway showing internal components.*

Modifications Required to Meet the Surface Threat

Detection and Tracking

Until recently, the eyes of Phalanx (in the form of its radar) were turned toward the skies, where radio frequency (RF) performance is perfectly adequate for detecting and tracking an incoming target. However, the surface of the water is far detached from that comfortable RF environment. The littoral environment close to land is not conducive to radar-controlled weapons, as the British Royal Navy discovered to its dismay in San Carlos' water during the 1982 Falklands War. Radar used in this environment suffers from the degrading effect of clutter.

Close to the surface, the restless sea causes clutter that is unwanted in returning echoes. The effect can be seen in Figure 3, which plots the returning signal strength for a typical radar working in free space, the amount of clutter return from sea state 5, and the combined effect if clutter were included as part of noise in the signal-to-noise (S/N) equation. For ranges out to 4 nmi, the returning signal would be below the detection threshold. When working at ranges over the land, it is totally swamped by the combined effect of noise and clutter. This area is precisely the littoral environment where Phalanx is now required to operate. Of course, this effect can be minimized by introducing moving-target indicator circuitry or doppler techniques, and Phalanx presently does this. However, the receiver processor was designed to look for a high-speed, supersonic missile target rather than slow-speed aircraft or a surface patrol boat masked by the sea clutter. Consequently, Phalanx would be unable to see and, in order to use RF efficiently in this environment, sophisticated radar equipment would be required, demanding a departure from the philosophy of NDI policy. It was, therefore, immediately evident that a different approach would be required.

The Introduction of Electro-Optics

Adopting an EO tracking device provides a solution to the RF constraint. EO devices can work regardless of the time of day. They are generally more accurate than their radar counterparts; and they are, of course, virtually undetectable and resistant to jamming. For a day-and-night capability against a target that is generally hotter than its surroundings, an infrared (IR) device can provide the capability required. With several devices already on the market, the NDI policy could be maintained. The question is: what sort of device are we looking for?

a. Active or Passive. With the target already providing the IR source (from engines, aerodynamic heating, or human body heat), active trackers can give way to passive tracking and, thus, reap the benefits of range advantage.

b. Range Finding. With the search radar available for use during EO-tracking, range could be determined by radar.

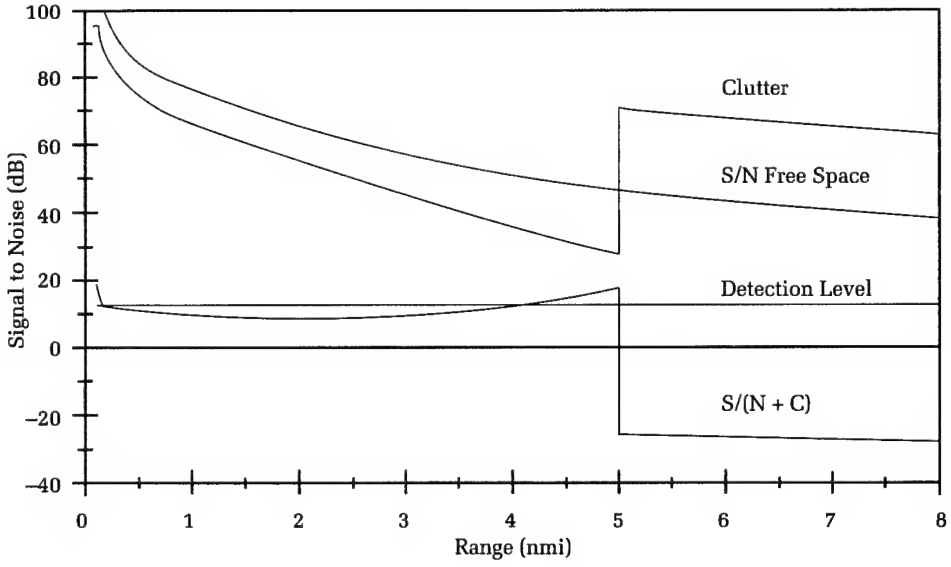


Figure 3. *The effect of clutter on a radar signal (sea state 5, land at 5 nmi range).*

Therefore, a laser range finder (given use of a passive detector) would not be required. However, as previously mentioned, RF performance alone would be unsatisfactory when sea state causes clutter. Thus, although the search radar will endeavor to be used for surface ranging, if this cannot be achieved, the operator must be able to manually input range, which can be determined from another radar set more appropriate to surface search. NSWCDD is already developing an automatic range input from sensor to Phalanx for a future upgrade. With the relatively slow speed of the target and the short firing range, NSWCDD has proved that the fire control of Phalanx will be able to tolerate the human input of range.

Wavelength

Current IR detection devices are available in the middle IR (MIR) at 3 to 5 μm and the far IR (FIR) at 8 to 12 μm. As it was known that changes in temperature and/or humidity could affect IR transmission, NSWCDD analyzed meteorological data from likely areas of U.S. Navy operations worldwide and demonstrated that, on average, IR transmission would be better in the MIR range. Further trials also measured the IR signatures of likely hostile targets to identify the most suitable operating wavelength for the surface-mode scenario. It was also found during a series of trials that 3 to 5 μm performance suffers from sun glint worse than the 8 to 12 μm range. This was determined to degrade the recognition ability of the system.

Effect of Smoke and Flash upon the IR Transmission

Despite the benefits associated with the sight working in the MIR, one other aspect IR hitherto never had to accommodate was the close proximity of the sensor to smoke and flash emanating from the ends of the gun barrels. The Dahlgren Division ran a series of trials to determine MIR and FIR transmission through the smoke and flash from a firing Phalanx gun.[1] The results are summarized in Figure 4 and show that, through this hostile transmission environment, the FIR performance can be considered to be better. All these previously described factors have been researched by the Division and presented to HMSC to enable them to select the most appropriate Forward-Looking Infrared Radar (FLIR). The result was that an 8 to 12 μm FLIR is being procured, as the performance in the presence of sun glint, and smoke and flash outweighed the longer range performance, which is outside Phalanx Surface Mode Requirements.

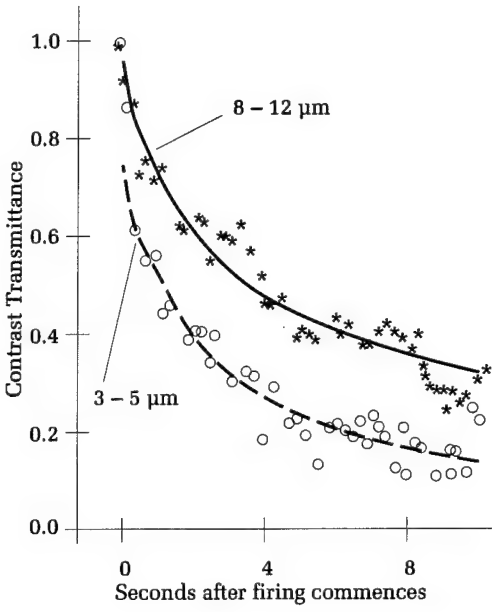


Figure 4. *IR transmission through smoke and flash of a firing Phalanx gun.*

Is Phalanx Effective in the Surface Mode?

With the EO sight determined, Phalanx can, at least, be pointed successfully at the target. However, will the gun system itself be effective? This was proven in a series of firing trials against a patrol boat representative of the surface-mode threat. It was assumed that only a few well-placed rounds would be required to disable such a craft—by either ensuring its mechanical disablement, detonating onboard ammunition, or even killing the crew! Determining the optimum number of rounds required was important, as Phalanx must remain able to complete its primary mission of engaging the antiship missile, and an unnecessary expenditure of bullets in a surface engagement could diminish this capability. The diagram in Figure 5 shows the extent of damage only a few rounds would make to a patrol boat. It is unlikely the boat would be sunk, but both the crew and the machinery would be severely damaged, and the boat disabled.

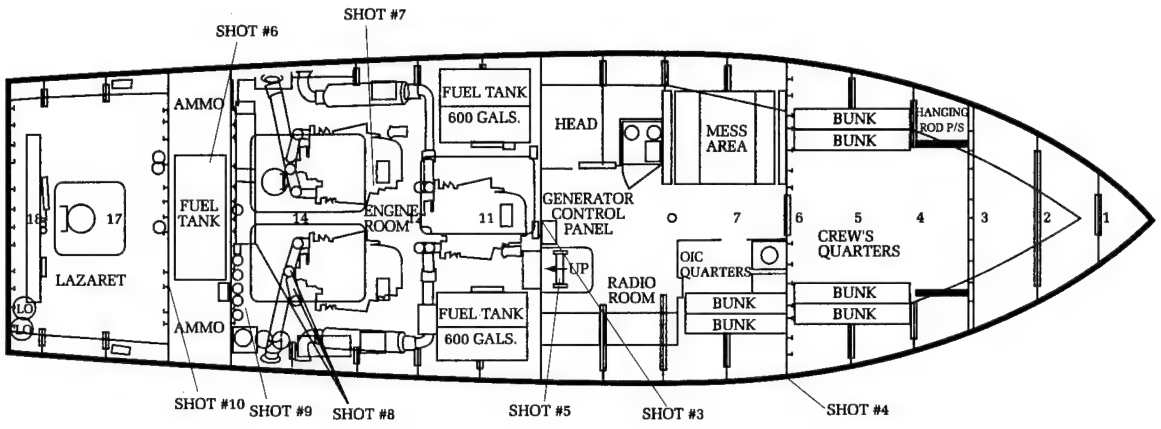


Figure 5. *Firing trials against a patrol boat. (Mk 149 Armor-Piercing Discarding Sabot (APDS) empirical shot lines against 65-foot Mk 3 PB777 at 1550-meter range.)*

Can the Primary Mission Still Be Achieved?

Program requirements specify that Phalanx must retain its primary mode of antiship missile defense. Since tracking the surface-mode target is now accomplished through the EO sight, the search radar can still actively scan for any incoming threat. These air targets will be evaluated by the Phalanx onboard computer and, if determined to be a threat, the system will break out of the surface mode to engage the incoming target. Figure 6 shows that, even with a surface-mode engagement in process, an incoming missile can be engaged and destroyed, and the surface target re-engaged to destruction, all within the minimum keep-out range and within one minute. The mission will, therefore, be achieved.

Improvement to Phalanx Primary Mission from Surface-Mode Introduction

The original requirement that Phalanx retain its primary mission of antiair warfare (AAW) has indeed been demonstrated. Moreover, EO sensor trials conducted in 1991 revealed that incorporation of the EO sensor required for surface-mode tracking enhances the AAW-mode tracking.

Joint NSWCDD/HMSC tracking trials involved the simultaneous use of both radar and EO while recording the quality of track. Figure 7 summarizes the results in graphs, depicting the quality of track from an EO sight and a radar when tracking the same target. The EO sight clearly gives a smoother track. This is an important point when remembering that the incoming missile is

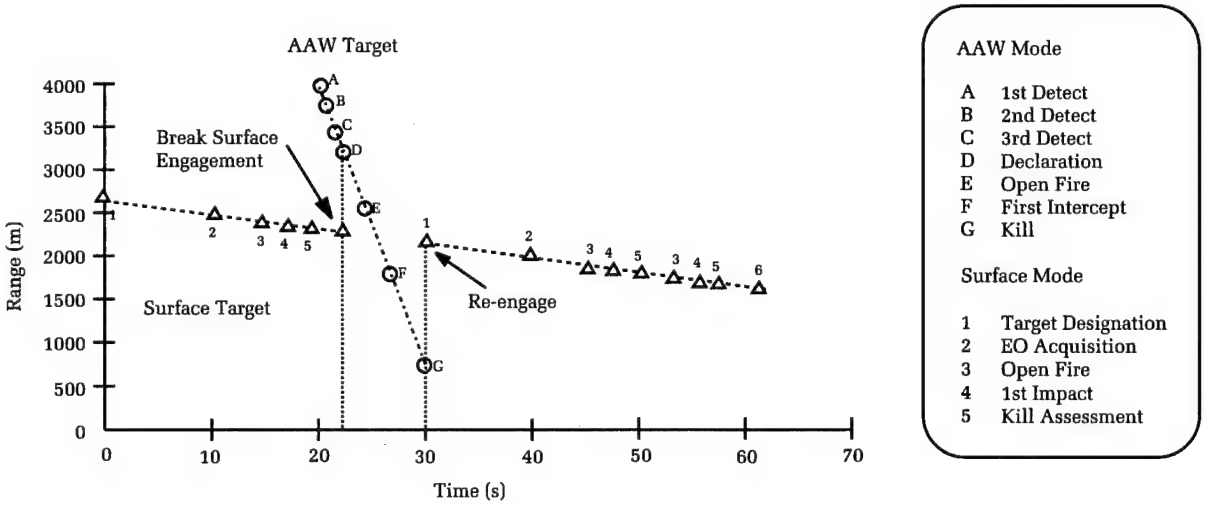


Figure 6. *Events of a surface and AAW engagement.*

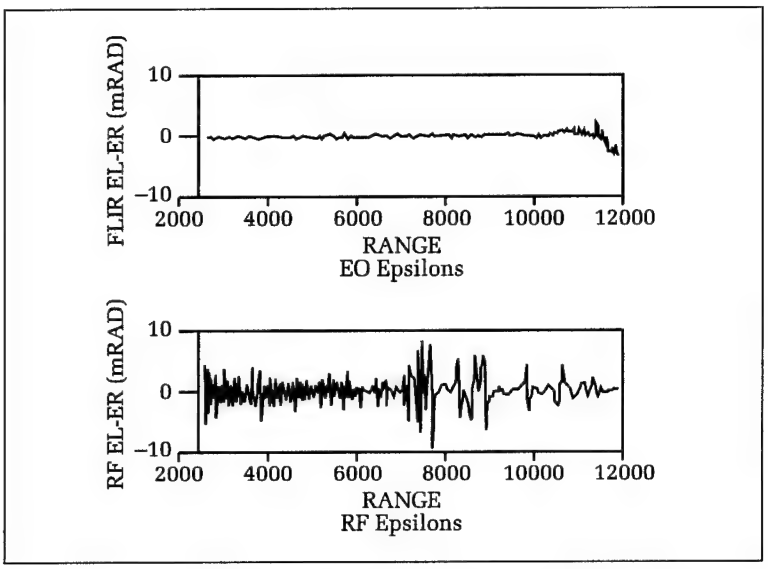


Figure 7. *Quality of track from EO and radar sensors.*

about 0.5 m in diameter and that the gun is aimed by the tracker. Any deviation of this point-of-aim results in bullets lost. Switching to EO improves the quality of track, increases the number of hits on the target and, thus, the probability of a kill. (Gun dispersion will be the same for both EO and RF tracking.) Integration of the EO track facility to AAW engagements will, therefore, be incorporated into later derivatives of Phalanx (Baseline 3).

Surface-Mode Integration to Phalanx

Minor equipment modification to the existing Phalanx system will be required to enable surface target engagement. In addition to the EO sight, a separate tracker, purely for surface-mode tracking and fire-control generation (again an NDI item), will be incorporated. As the surface-mode mission requires the system to operate in the presence of friendly and hostile craft, surface-mode target acquisition is under the control of an operator who can visually recognize the target before engagement and complete the acquisition process—an otherwise complex achievement. This capability is limited to surface targets, when time scales are significantly longer than in the AAW environment.

Program Status

The program is now far advanced and passed the System Design Review, where the Department of Defense endorsed the final concepts. A prototype is expected mid 1995 for proving firings and further development. Introduction to the fleet is planned to begin in 1997/98.

Conclusions

The recent change in naval tactics has led the U.S. Navy into the area of littoral warfare and towards a new threat: the fast-moving surface patrol craft, which could be engaged by an advanced machine gun. It was concluded that the Phalanx system, already universally fitted throughout the fleet, could be modified to counter this threat, and the concept of introducing a Phalanx surface mode was raised. In order to minimize costs and to expedite fleet introduction, maximum use is being made of NDI.

Perhaps the most radical departure from traditional techniques is the incorporation of EO IR sights. Addition of the EO will enable target tracking close to the water in an environment hostile to RF operations. It will also improve the tracking quality of AAW operations and increase the probability of kill against inbound antiship missiles. It has been demonstrated that the present M61A1 gun, firing the Mk 149 round, can inflict crippling damage to both the material and crew of a hostile patrol boat, thus defeating the threat. At the same time, Phalanx will remain ready to revert to its primary mission of engaging the incoming antiship missile before returning to the surface target.

NSWCDD, as the TDA throughout this development, has accrued the many benefits of incorporating NDIs into a mature system while, at the same time, achieving a significant performance upgrade. While proving that the available equipment could deliver the performance, the Division also determined that the existing components (such as the gun itself) could meet the new requirements. Therefore, NSWCDD has been pivotal in achieving a significant increase to fleet capability in the most timely and cost-effective manner.

Reference

1. Austin, D. E., Carr, R. B. and Wardlaw, M. J., *Test Results of Smoke/Flash IR Effects for the Phalanx Close-In Weapon System,* NAVSWC TR 91-587, Naval Surface Warfare Center, Dahlgren, VA.

The Author

LIEUTENANT COMMANDER STUART BORLAND joined the British Royal Navy in 1984 after having read a degree in engineering at Coventry University in England. He joined the Royal Navy as a specialist officer in the Weapon Engineering Branch, which is responsible for the maintenance of all weapon, computer, and sensor systems on board Royal Navy warships. He joined the Britannia Royal Naval College at Dartmouth (previous students include Princes Charles and Andrew) and completed general training deployments aboard HMS FEARLESS (LPD) and HMS NOTTINGHAM (DDG). He undertook specialist training at the Royal Naval Engineering College in Plymouth and HMS COLLINGWOOD in Portsmouth before joining HMS EDINBURGH (DDG) as the Weapons System Officer, where he was decorated for service in the Gulf in 1987. After a two-year tour aboard HMS EDINBURGH, he was appointed to the UK Phalanx procurement program as the technical advisor on all Phalanx matters. During this appointment, he paid several visits to the U.S. and NSWCDD. From 1992 to 1993, he served on the staff of the Admiral of the Surface Flotilla. Following promotion to Lieutenant Commander, he was appointed as the exchange officer to NSWCDD in November 1993, where he assists the Phalanx program manager in system engineering. He became a Chartered Engineer and was invited to join the prestigious Institute of Electrical Engineers in 1992.

Nonlinear Least-Squares Estimation in Naval Gun Fire Control

Donald de Forest Boyer, Edward L. Price, and Douglas O. Haas

This article describes a concept for improving the accuracy of naval gunnery. The key to this concept is a procedure that estimates a set of parameters used in solving the ballistics problem. The basic idea is to track a projectile as it is fired, then produce a set of parameters that allows the mathematical trajectory (used in solving the ballistics problem) to closely match the measured trajectory. This process differs from the time-honored method of spotting the fall of shot, in that it tunes the model used in generating the gun orders, rather than simply moving the impact point.

The key to the procedure is a Finite Difference Gauss-Newton technique, that processes the track data and extracts estimates of muzzle velocity, meteorological data, and alignment errors. These estimates are then combined and used in the fire-control processing to compute a solution of increased accuracy relative to the conventional methods.

From the initial concept formulation, through tactical software development and implementation aboard the Arleigh Burke class of AEGIS destroyers, the work has been performed entirely by the Naval Surface Warfare Center Dahlgren Division (NSWCDD).

Introduction

Because this article deals with one aspect of gun fire control, it will be helpful in placing this particular aspect in the proper perspective to briefly review the general problem to which gun fire control is devoted. A functional diagram of a typical naval gun fire-control system is presented in Figure 1. Position measurements of a target are provided by a tracking sensor and input to the fire-control computer. The measurements are put in a convenient reference frame by a sequence of transformations that translate to the desired reference origin and rotate to compensate for angular displacement of the ship. The stabilized target-position measurement data are processed by a track-filter algorithm that provides estimates of target position and motion (antiair and surface direct-fire missions). These target-state estimates are used by the target-trajectory extrapolation algorithm to provide estimates of target position at any specified future time. The gun-command function computes pointing angles for the gun that cause the computed trajectory of the projectile to intercept the computed trajectory of the target. The accuracy of the process depends on how accurately projectile and target flight are modeled in the computer.

For some years now, naval gun fire-control systems have modeled projectile flight by numerically integrating a system of nonlinear differential equations of

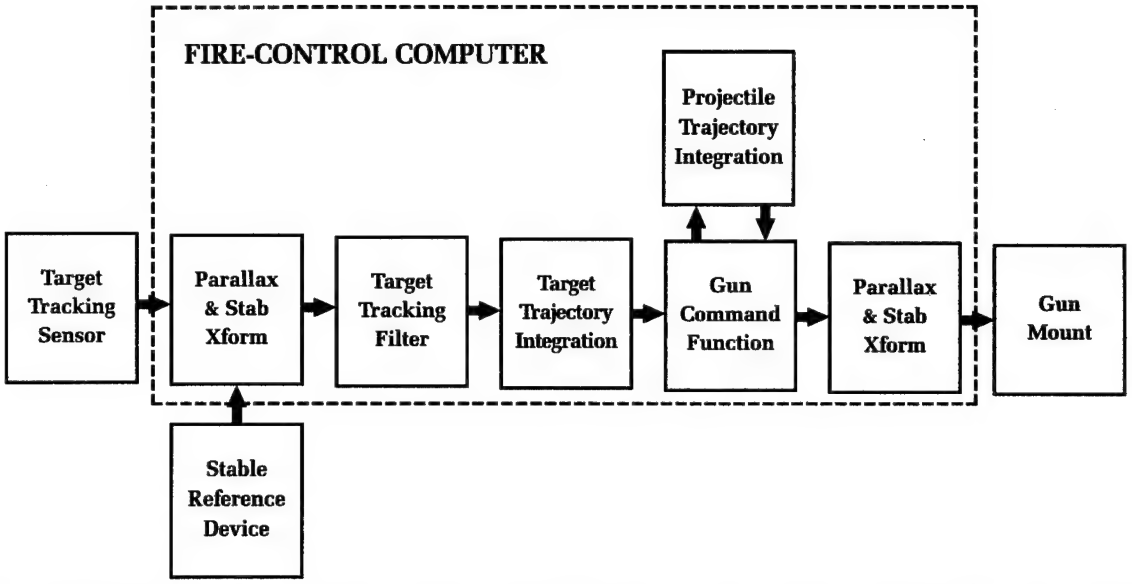


Figure 1. *Typical naval gun fire-control system with no corrective feedback to address bias errors.*

motion, which include the Coriolis effect and drift for spinning projectiles. For a thorough treatment of ballistics, see *Exterior Ballistics.*[1] In order to use the projectile model in the iterative procedure described above to compute the gun-pointing angles and time-of-flight, it is necessary to have accurate estimates of the initial conditions and ballistics parameters. The parameters under consideration fall naturally into two categories: (1) initial conditions and (2) meteorological (or atmospheric) conditions. Included in the initial conditions are initial projectile speed (muzzle velocity) and the gun-pointing angle at the time the projectile is fired.

Data providing nominal muzzle velocity as a function of the type of gun, projectile, fuze, charge, powder temperature, and barrel wear have been developed and are in general use. Experiments have shown, however, that in spite of these considerations there is typically a mean error in the muzzle velocity that is significant, and that this mean error is typically greater than the deviation from shot to shot. Regarding the gun-pointing angle at time-of-fire, the pointing command is known, of course. Uncertainty in the actual pointing angle arises due to alignment errors between the tracking sensor and the gun mount. Misalignment may be in the form of drive axis tilt or in differences between indicated train and/or elevation angle and the actual values of these angles.

The conventional approach to dealing with misalignment in naval shipboard gun systems is to perform a painstaking mechanical and electrical alignment procedure. This procedure is time-consuming and difficult to accomplish at sea. Even when an accurate system alignment is achieved, the condition is temporary. Numerous effects such as temperature changes, changes in ship loading, other sources of mechanical stress, and varying component characteristics due to aging contribute to alignment deterioration. This is one of the most serious contributors to gun-fire error in surface direct and naval gun fire-support missions and is second only to prediction of target future position in contributing to error in antiair engagements.

The standard source of data on atmospheric conditions available for use in naval operations is the Met Message. This message gives a tabulation of vertical profiles for wind speed and direction, atmospheric density, and temperature. The Met Message is compiled at one of numerous meteorological facilities located around the globe. It is generally unsatisfactory as a source of information for gun fire control for two reasons: (1) the recording station is usually located far (frequently hundreds of miles) from the ship engaged in the gun firing exercise, and (2) the data may be excessively stale.

Coriolis Effect: The apparent deflection of projectile motion that is a result of observing the flight from a noninertial, uniformly rotating frame of reference, such as earth.

Drift of Spinning Projectiles: The lateral deviation of the point of fall from the original plane of fire, due only to effects set up by the rotation of the projectile.

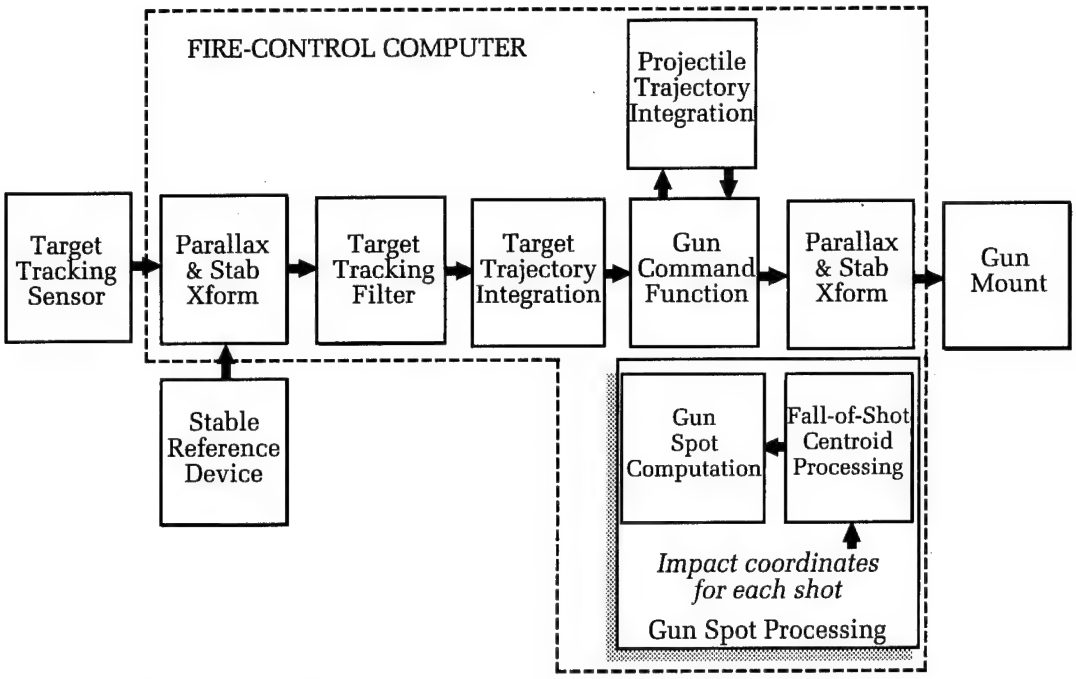


Figure 2. *Conventional spotting to correct for biases in gunfire.*

The spotting procedure is a common technique for improving the accuracy of gunfire against surface targets that has been in use since the earliest development of gun technology. Figure 2 illustrates how spotting is incorporated in the fire-control process. In a typical application of this procedure, a number of shots are fired at a specific target. The centroid of the fall of shot is computed, and incremental adjustments to the gun-pointing angle are made to move the centroid onto the target. While this technique is effective under certain conditions and will continue to have a role in the future, it has some fundamental limitations. Spotting is a means to compensate for the combined effects of all sources contributing to error in projectile delivery, but is incapable of identifying the individual culprits. Because the influence of some errors is dependent on direction of fire, range of target, and other aspects of the relation of gun to target, spotting is effective in improving accuracy only for a target in the vicinity of that used in the spotting exercise. Another limitation is that basic instabilities make it difficult to achieve progressive accuracy improvements by making shot-by-shot adjustments in gun pointing-angle corrections. Thus, in the interest of statistical stability, shots are processed in batches to develop the spots used in this process.

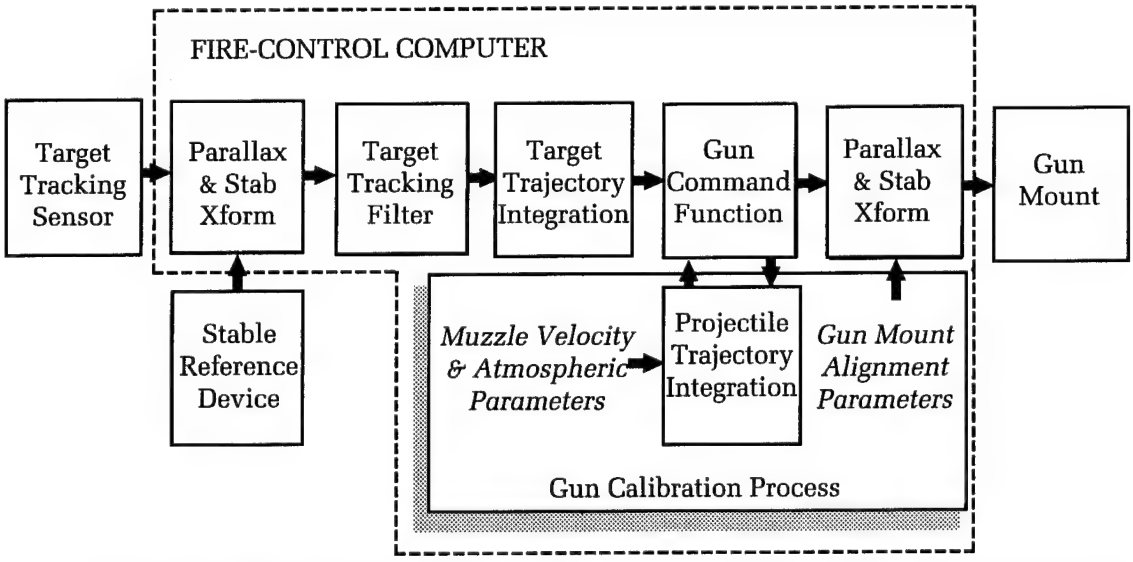


Figure 3. *Gun calibration using projectile track data to address gunfire errors at the source.*

The preceding discussion has emphasized that inaccurate information on initial conditions of projectile motion and on atmospheric conditions substantially limits the accuracy of naval gunfire, and that conventional means for addressing this limitation are inadequate. In the next section, a technique to obtain accurate estimates of these parameters is presented. The technique exploits the capability of modern radar systems to track gun projectiles and provides a means to extract the parameter estimates from projectile track data, thereby significantly improving gunfire accuracy. The relationship of a calibration process based on this technique to the overall gun fire-control system is depicted in Figure 3.

Problem Statement

In general terms, the problem addressed is the development of a procedure to provide estimates of input parameters to a projectile trajectory (ballistics) algorithm that will cause the computed trajectory to coincide with the actual one. (It is assumed that an algorithm modeling projectile flight is available that represents actual flight with precision when accurate values of the input parameters are provided.) The parameters involved are muzzle velocity, meteorological data, and alignment errors. The following description of an operational application to estimate the parameters in question will help to motivate the mathematical development and to put it in perspective.

A gun projectile is fired and tracked by a radar system over part of its trajectory. The projectile track data are input to the fire-control computer, transformed to a Cartesian earth-fixed reference frame with origin at the gun mount at time-of-fire, and stored. This frame is illustrated in Figure 4. Using a computational model of projectile dynamics and an initial guess of the parameter set to be estimated, a trajectory is calculated to coincide in time and space with the observed segment. A systematic technique, depicted in Figure 5, is then employed to compute an adjustment to the parameter set that will move the next calculated trajectory into closer alignment with the measured trajectory. This process is iterated until the desired degree of convergence has been achieved. Now available are estimates of the parameter set that result in computed trajectories corresponding more closely to the actual ones, thereby providing greater accuracy in subsequent gunfire.

The improved parameter set is then subjected to further processing to produce a set of values that is in the proper format to be entered into the fire-control program. These values are then statistically combined with previous results. This process allows the estimates to adapt (in real time) to changing conditions and, through interpolation, be made applicable to a wide range of target ranges and gun bearings.

The preceding description is an oversimplification of the implementation in the Arleigh Burke class of destroyers, but serves to illustrate the main characteristics and to establish a useful context for the mathematical development to follow.

In order to put the qualitative description of the estimation procedure given above into a mathematical formulation, some preliminary definitions and notations are required. Let x be the position vector of the projectile in rectangular gun-centered coordinates, and let p be the vector of parameters to be estimated. The equations of motion for the projectile can be written

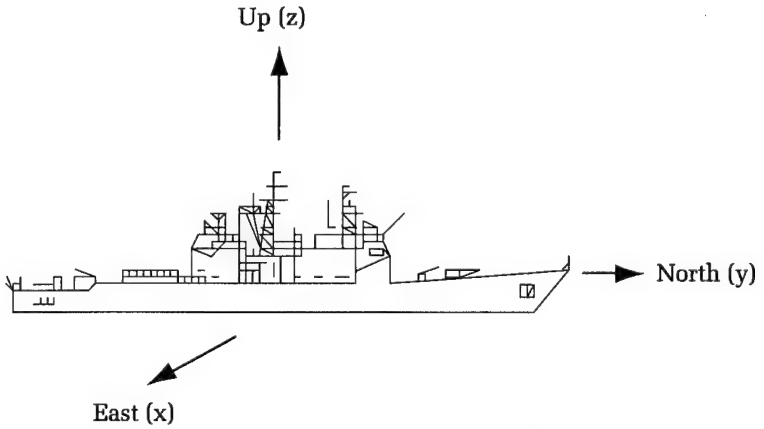


Figure 4. *The basic coordinate reference frame associated with the problem statement and analysis is ship-referenced, earth-fixed Cartesian.*

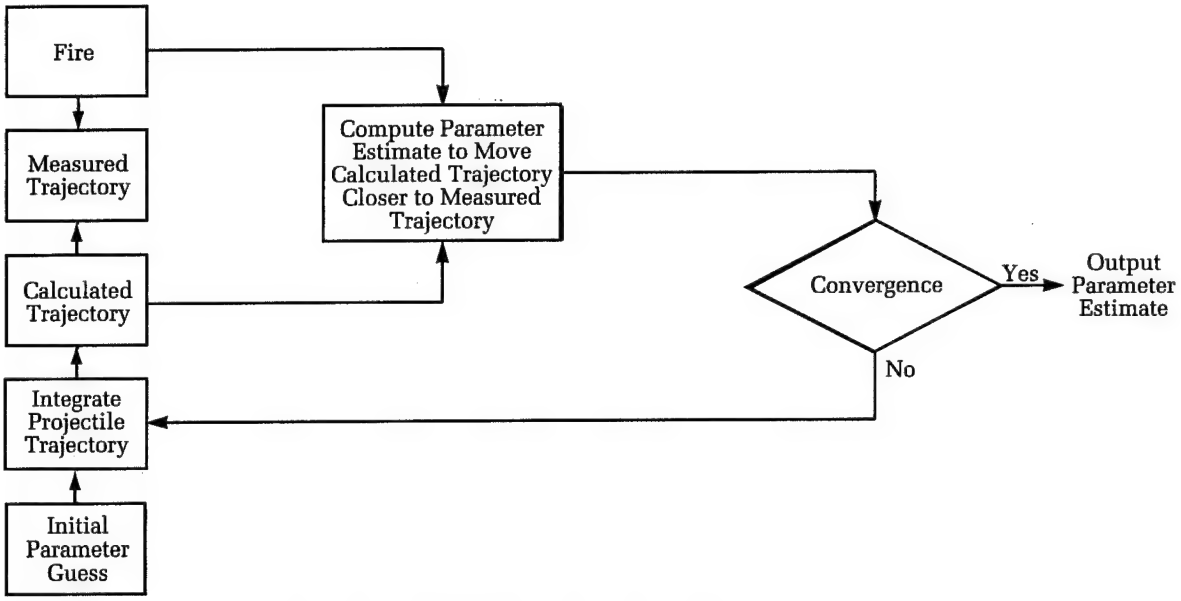


Figure 5. *Calibration parameter estimation process.*

$$\begin{aligned} \ddot{x}(p,t) &= f(x,\dot{x},p,t) \\ x(p,0) &= 0 \\ \dot{x}(p,0) &= v_0(p) \end{aligned} \qquad (1)$$

Now, let us assume that the radar records measurements of the projectile at times (t1,...,tn), and let

$$Z = (Z_1, Z_2, \ldots, Z_n)^T$$

be measurements from the radar; note: $Z_i=(x_i,y_i,z_i)$, so that there are $3n$ components in the measurement vector. Now we also discretize the numerical solution to the differential Equations (1) in the same manner

$$X(p) = (x(p,t_1), x(p,t_2), \ldots, x(p,t_n))^T$$

so that there are $3n$ components in the discretized trajectory model. Now we write

$$Z = X(p) + V \qquad (2)$$

where V is an error vector that includes sensor errors as well as modeling errors. If the covariance matrix $R = E(V \cdot V^T)$ is readily available, then the standard technique used to obtain statistically meaningful estimates of the parameter vector, p, is to find a vector, p, that minimizes the function

$$F(p) = (X(p) - Z)^T R^{-1} (X(p) - Z) \qquad (3)$$

Applying the above minimization procedure generates the minimum variance, or Gauss-Markov estimate, of the parameter vector in question. For a thorough discussion of estimation theory, see *Parameter Estimation.*[2] Through extensive testing with both simulated track data and real track data, it was determined that suitable values for the ballistics parameters could be obtained by letting R = I, the identity matrix. So the problem was formulated so as to minimize the objective function

$$F(p) = \frac{1}{2} (X(p) - Z)^T (X(p) - Z) \qquad (4)$$

and find the nonlinear least-squares estimate of the parameters.

A Finite Difference Gauss-Newton Method

In order to find the minimum value of the above objective function, it will be necessary to compute the gradient of $F(p)$

$$\nabla F(p) = J(p)^T (X(P) - Z) \qquad (5)$$

where

$$J(p) = \begin{pmatrix} \frac{\partial X_1}{\partial p_1} & \cdots & \frac{\partial X_1}{\partial p_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial X_n}{\partial p_1} & \cdots & \frac{\partial X_n}{\partial p_m} \end{pmatrix}$$

The vector-valued function of the ballistics parameters, $X(p)$, that appears in the objective function in Equation (4), resulted from the discretization of a numerical solution of the system of nonlinear differential Equations (1). Since this is admittedly a rather obscure manner in which to define a function, it is not surprising that closed-form expressions for the partial derivatives of the trajectory function with respect to the parameters are not readily available. Thus, when we refer to partial derivatives in this article, we mean finite

differences computed as follows

$$\frac{\partial X_i}{\partial p_j} = \frac{X_i(p+\varepsilon_j U_j) - X_i(p)}{\varepsilon_j}$$

where U_j is the j-th unit vector in m dimensional Euclidean space,

and

$$\varepsilon_j = \begin{cases} \|p_j\|_2 \bullet ep & \text{if } \|p_j\|_2 \geq 1 \\ ep & \text{if } \|p_j\|_2 < 1 \end{cases}$$

$$\varepsilon p = \sqrt{machine\ epsilon}$$

and

So if one implements these finite differences on a machine that uses hexadecimal arithmetic with single precision, then $ep = \sqrt{16^{-5}} \approx 10^{-3}$. The above choice of differencing parameters makes sense, and the parameters to be estimated can vary significantly in magnitude; and we certainly do not wish to allow the differencing parameters to be too small. For a careful discussion of these topics, see *Practical Optimization*, pages 127–133 and 339–345.[3]

A necessary condition that a smooth function must satisfy if it is to possess a relative minimum at an interior point, p_0 of a set, is that $\nabla F(p_0) = 0$. Thus, it is traditional to search for local minima among the stationary points. In order to even consider the task of setting the gradient in this problem equal to zero, with the faint hope of ever attempting to solve the resulting system of equations, the trajectory function, $X(p)$, must be linearized. So we proceed and let

$$X(p) = X(p_0) + J(p_0)(p - p_0) \qquad (6)$$

This replaces the nonlinear function $X(p) - X(p_0)$ with a linear function, which linearizes the nonlinear function about the trajectory $X(p_0)$. Replacing $X(p)$ in the gradient expression results in (to simplify the formula, let $J_k = J(p_k)$ and similarly let $X_k = X(p_k)$

$$\nabla F(p) = J_0^T[X_0 + J_0(p - p_0) - Z] \qquad (7)$$

By setting the gradient equal to 0 and rearranging the matrices, we have

$$J_0^T J_0(p - p_0) = -J_0^T(X_0 - Z)$$

We can now solve explicitly for $\hat{p}$

$$\hat{p} = p_0 - (J_0^T J_0)^{-1} J_0^T(X_0 - Z) \qquad (8)$$

Now if $X(p) - X(p_0)$ were actually equal to the linear function on the right side of Equation (6), then $\nabla F(\hat{p})$ would be equal to 0, and $\hat{p}$ would be among the stationary points of the nonlinear least-squares objective function, F. Also in the case of a linear trajectory model, the matrix $J_0^T J_0$ would be the Hessian of $F(p)$; and because J_0 should always have full column rank, $J_0^T J_0$ will be positive definite and, thus, satisfy the sufficient condition that ensures that $\hat{p}$ is the least-squares estimate of p.

As the trajectory model in the gun-calibration problem under discussion is certainly a nonlinear function of the ballistics parameters to be estimated, an iterative procedure is necessary. Iterative procedures for solving nonlinear optimization problems generally produce a sequence of estimates $\{\hat{p}_1,...,\hat{p}_k\}$, such that for each i, $F(\hat{p}_{i+1}) < F(\hat{p}_i)$; they are appropriately called *descent methods*. A vector d_i in the parameter space is called a *descent direction* at p_i, provided one is able to find a scalar α_i (called the step length), such that $F(p_i + \alpha_i d_i) < F(p_i)$. The steps involved in building an iteration procedure that will find a point p_{i+1}, given the point, p_i, are:

(1) Determine a descent direction, d_i
(2) Compute the step length, α_i
(3) Set $p_{i+1} = p_i + \alpha_i d_i$.

It can be shown that if F is a smooth function, and A is a positive definite matrix, then $d_i = -A \cdot \nabla F$ is a descent direction. The expression for $\hat{p} - p_0$ given in Equation (8) is a descent direction at p_0, provided J_0 is the true Jacobian matrix at p_0 and has full column rank. The step given by Equation (8) is called the Gauss-Newton step if the Jacobian is exact, and it is called a Finite Difference Gauss-Newton step when the Jacobian is approximated, as indicated above. The iteration procedure that was actually implemented in the gun-calibration procedure was a Finite Difference Gauss-Newton Method

$$p_{k+1} = p_k - (J_k^T J_k)^{-1} J_k^T(X_k - Z) \qquad (9)$$

The step length is taken to be $\alpha_k = 1$ in Equation (9), where in a general iterative procedure for a nonlinear minimization problem, it is necessary to search for an α_k in order to ensure the descent condition. Computational experience has shown that nonlinear least-squares problems usually do not require a step length algorithm, even when approximate gradients and Hessians are employed. (See Chapter 10 in Dennis and Schnabel's book.[4]) The iteration formula given in

Equation (9) does not guarantee convergence for all nonlinear least-squares problems, but it does work very well in practice, especially when residual term $R(p) = (X(p) - Z)$ becomes small as p approaches the minimum value $\hat{p}$, and the initial value of the parameter vector p_0 is not "too bad." The original algorithm that was developed for this project was based on the Gauss-Markov estimate implemented computationally, with a finite difference Levenberg-Marquardt procedure.[4] This method, which is much more difficult to implement, guarantees convergence even in the large residual case and with no restriction on the starting point. However, it was determined, using very detailed radar and ship-motion simulations, that neither the statistical sophistication of the Gauss-Markov estimate, nor the computational power of the Levenberg-Marquardt algorithm is needed when applied to the gun-calibration problem. This is not surprising since, in this application, reasonable initial values for the ballistics parameters are available, and the least-squares formulation has reasonably small residuals. Thus, the Finite Difference Gauss-Newton iteration given in Equation (9) performs very well. In order to avoid the possible ill-conditioning of matrix inversion operation in Equation (9), QR-Factorization was used to solve

$$(J_k^T J_k)(p_{k+1} - p_k) = -J_k^T (X_k - Z) \quad (10)$$

The matrix $J_k^T J_k$ is written as a product of an orthogonal matrix, Q, times an upper triangular matrix, R, so Equation (10) becomes

$$QR(p_{k+1} - p_k) = -J_k^T (X_k - Z)$$

Since Q is orthogonal

$$R(p_{k+1} - p_k) = -Q^T J_k^T (X_k - Z)$$

Since R is upper triangular, the equation can be solved using back-substitution; this simplifies and stabilizes the computation significantly.

Secondary Processing

The foregoing presentation has described a technique for extracting certain parameters used in gun-ballistics computations by processing track data from projectiles fired by the gun. In an operational implementation, practical considerations require additional processing of the set of parameter estimates derived from each individual shot in order to get the most benefit in improved accuracy for subsequent shots. This additional, or secondary, processing serves two purposes: first, it puts the data in the proper format for use in ballistics processing and, second, it takes advantage of data from multiple shots to refine the accuracy of the parameter estimates. The data resulting from secondary processing are classified into three categories:

- meteorological data
- alignment data
- projectile initial velocity

The secondary processing is described in more detail in the remainder of this section.

Met Message Processing

The meteorological data consists of ballistic wind speed and direction and ballistic density organized in the form of a "Met Message." As mentioned in the Introduction, this is not considered to be a satisfactory source of information due to the distance between the measuring station and the ship, and the potential "staleness" of the data. The process described below is equivalent to measuring the data directly from the ship.

The Met Message is organized into 16 altitude zones corresponding to the apogee of the bullet trajectory in question, with ballistic values of wind speed and direction, and density. "Ballistic" values of these parameters are computed such that if single values are used in the process of integrating the bullet trajectory through its entire course of flight, the terminal point of the trajectory will be the same as if the exact values for the parameters (including variation with altitude) were used in the integration process. The following process was designed so that meteorological data in the Met Message format could be produced as part of the calibration procedure under discussion without needing to fire an exorbitant number of projectiles in the calibration exercise. This process involves partitioning the altitude coordinate into four zones. In the calibration exercise, the apogee, wind speed and direction, and density determined for each shot processed are recorded. Data for the individual shots are associated with one of the four zones according to the apogee of the projectile. When all shots are fired for a particular exercise, a linear-regression procedure is used to fit a curve to each set of parameter data (representing variation of that parameter with respect to altitude within the zone addressed) in each of the four zones. The set of curves resulting from this regression procedure is then used to fill in a Met Message table in the standard format.

Alignment Processing

Two types of errors are addressed in alignment processing:

- biases in gun train and elevation relative to commanded (or input) values
- tilt of the gun-train axis relative to deck normal (this is frequently described as tilt of the gun-train bearing roller path plane relative to deck plane)

The basic calibration procedure, in processing track data from a single shot, produces estimates of error in gun-pointing angles in the form of error in bearing and elevation in a stable (earth-fixed) reference frame. The alignment-error model described above includes four parameters: two each for drive response bias and for train axis tilt. Hence, at least two shots, with different relative bearing angles, are required to compute the full set of four alignment-error parameters.

A linear relationship among the four alignment-error parameters and the gun-pointing errors, as determined by the basic calibration procedure, has been developed. This relationship provides a means for estimating the pointing bias and roller-path tilt parameters (the four alignment-error parameters); e.g., using a least-squares fit to the gun bearing and elevation data collected from the set of shots processed in the calibration exercise.

Initial Velocity Processing

Secondary processing to refine the estimate of projectile initial velocity consists of executing a simple average of the velocity measurements for each individual shot over the set of shots fired for each fuze/projectile/charge combination used in the calibration exercise.

Implementation

The gun-system calibration procedure described in the preceding portion of this article is currently implemented on the Arleigh Burke class AEGIS destroyers. The procedure is incorporated in the Gun Calibration Program (GCP) resident in the Mk 160 Mod 4 Gun Fire-Control System (GFCS) that is part of the Mk 34 Gun Weapon System on the Arleigh Burke class. A diagram of this system highlighting the gun-calibration process is presented in Figure 6. Target track capability is provided to the GFCS primarily by the AN/SPY-1D radar system, with secondary support from an optical sight. The computational resources of the Mk 160 Mod 4 GFCS comprise two AN/UYK-44 computers, one of which maintains the interface with the rest of the AEGIS Combat System, while the other controls the gun mount.

The GCP is coded almost entirely in a high-level language (CMS-2); assembly language is used only where absolutely required. As the GCP is a computationally intensive process requiring that 10 projectile trajectories be integrated for each round processed in the calibration procedure, it is run as a background task. Typical execution requires 7 to 10 seconds of elapsed time.

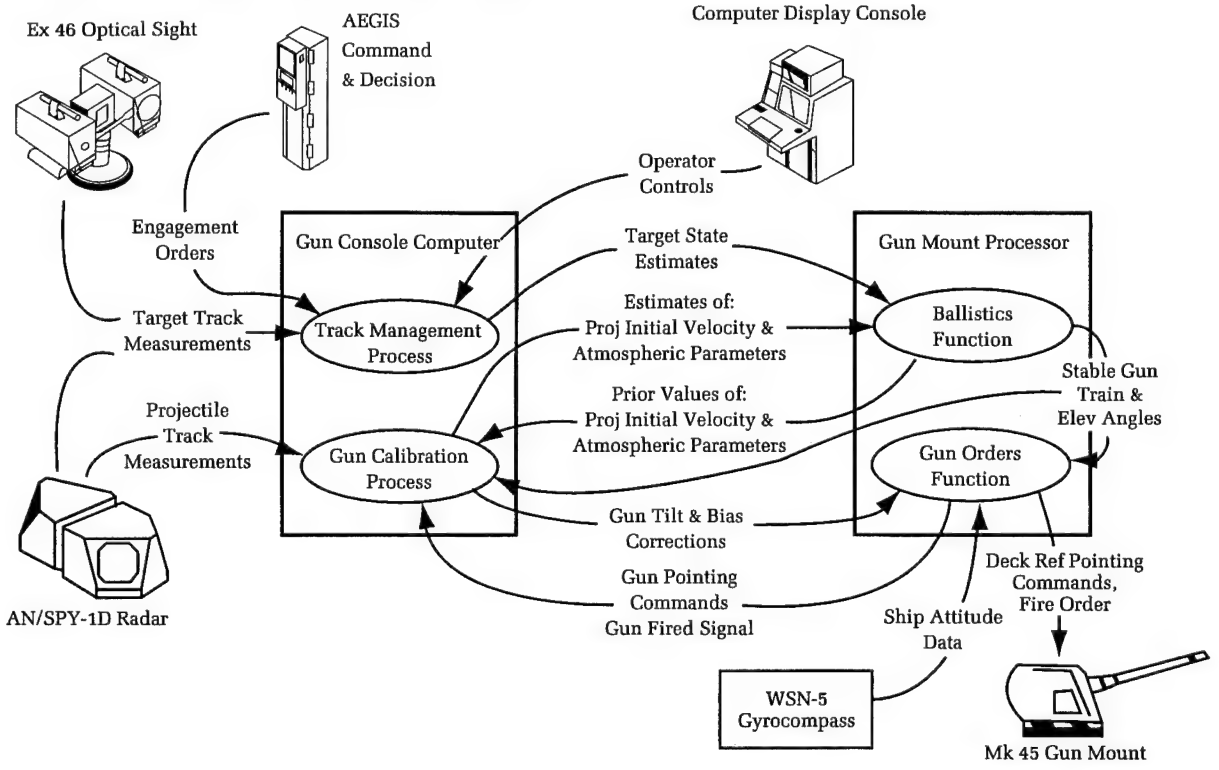


Figure 6. *The gun fire-control system on the USS Arleigh Burke (DDG 51).*

Figure 5 describes the basic calibration procedure as discussed in the Problem Statement. The measured trajectory is created by transforming the projectile track data to a stable frame, with origin at the gun mount at time-of-fire. An initial parameter set is used to compute a projectile trajectory that is stored in the form of position values corresponding to each time point for which a measurement exists in the measured trajectory. The Gauss-Newton procedure described in "A Finite Difference Gauss-Newton Method" is then applied to produce a new parameter set that will bring the calculated trajectory closer to the measured one. The new parameter set is then used to produce a new calculated trajectory to be employed in the next pass through the Gauss-Newton procedure. Two passes have been found to be adequate to achieve convergence in the Arleigh Burke implementation.

Figure 7 illustrates the secondary processing as implemented. Inputs are indicated both from the basic (Gauss-Newton) procedure and from fire-control processing associated with each shot fired in the calibration exercise. The outputs are produced as described in Secondary Processing.

Performance Results

Convergence of the Gauss-Newton procedure is illustrated in Figure 8. Shown in this figure are the measured trajectory, the reference trajectory (the trajectory produced from the initial parameter set), the trajectory produced from the parameter set obtained on the first pass through the procedure, and the trajectory produced from the parameter set obtained on the second (and final) pass. The tendency to bring the computed trajectory into alignment with the measured one is evident. It is interesting to observe that, in the absence of the calibration procedure, the trajectories produced by the initial parameter set would be used in the process of computing gun orders as part of normal gun fire control. An indication of the improvement in gun-pointing accuracy is given by comparing trajectories produced from the initial and final parameter sets with the measured trajectory.

Another demonstration of the convergence behavior of the Gauss-Newton procedure is presented in Table 1. Here, the root-sum-square of the residual difference between points on the measured trajectory with each of the computed trajectories described in the preceding paragraph is shown for trajectories at three different ranges. The degree of convergence is dramatic.

Estimation of alignment parameters in the calibration process is illustrated in Table 2. In this particular exercise, the parameters indicating roller-path tilt were set to zero in the initial parameter set. Actual tilt was 0.07 degree, with a high point at a train angle of 237 degrees (this is the format for roller-path tilt in the Mk 160 Mod 4 GFCS implementation) as shown in the table. The data indicates stable convergence with very good accuracy obtained after processing 15 rounds.

Finally, Figure 9 shows how the density portion of the Met Message is constructed from data collected from individual shots in the calibration process. The figure shows a plot of estimated density values for a number of shots, with apogee falling within band 2 of the four-band partition of altitude. Also indicated is the linear least-squares fit to these data points from which entries in the Met Message table are obtained. The solid vertical lines extending below the linear curve correspond to altitude values in the center of

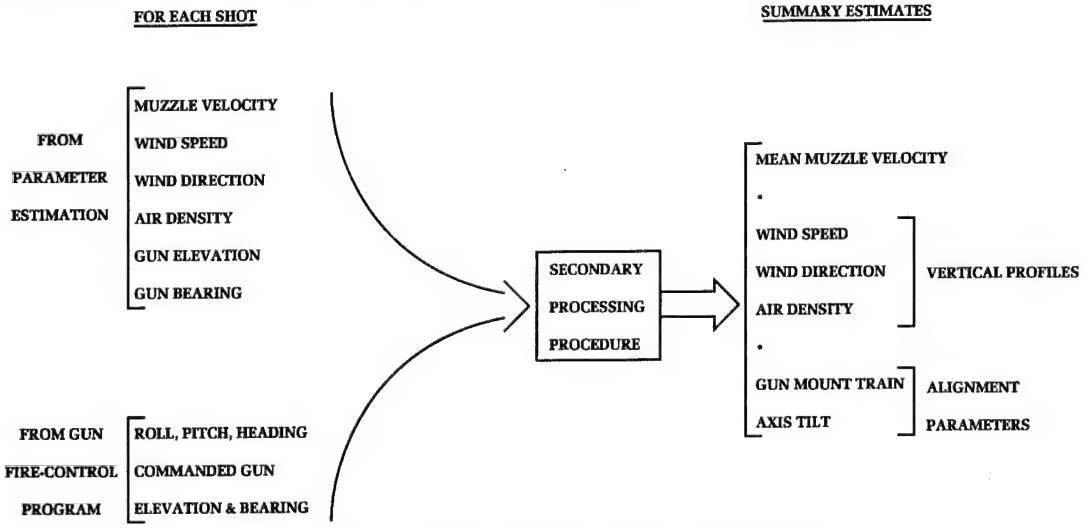


Figure 7. *Calibration secondary processing.*

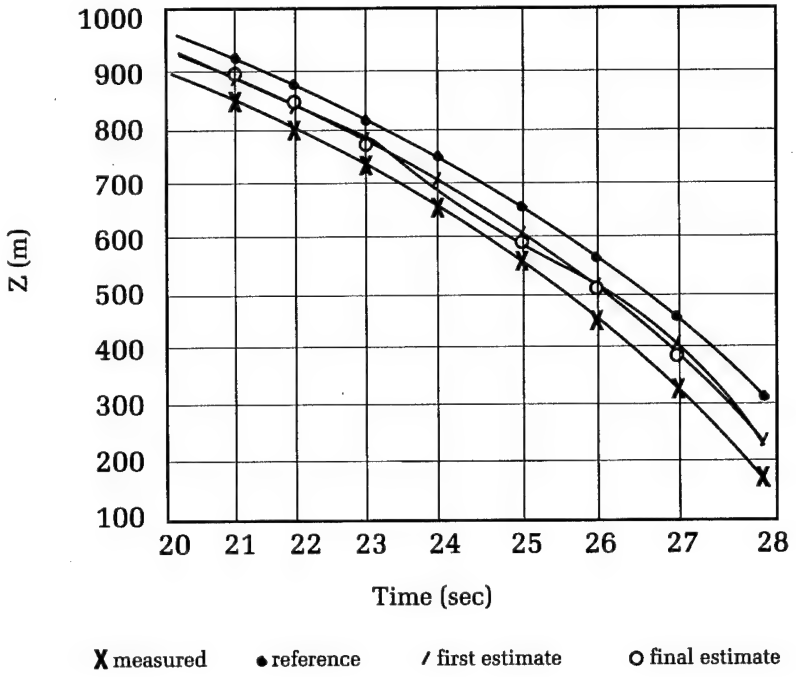


Figure 8. *Trajectories using various parameter estimates.*

Table 1. *Residual Analysis at Various Ranges (in Yards)*

Range	Reference	First Estimate	Final Estimate
11000	34597.60	108.67	105.77
19000	838134.23	118.57	75.65
24000	1709533.42	3525.39	65.17

Table 2. *Alignment Results*

Round	Inclination (Deg)	Bearing (Deg)
1	0.037643	242.39
2	0.069849	251.17
3	0.070883	245.80
4	0.072289	239.58
5	0.072429	237.80
9	0.070411	238.28
10	0.079528	238.60
15	0.071737	238.24
TRUE	0.070000	237.00

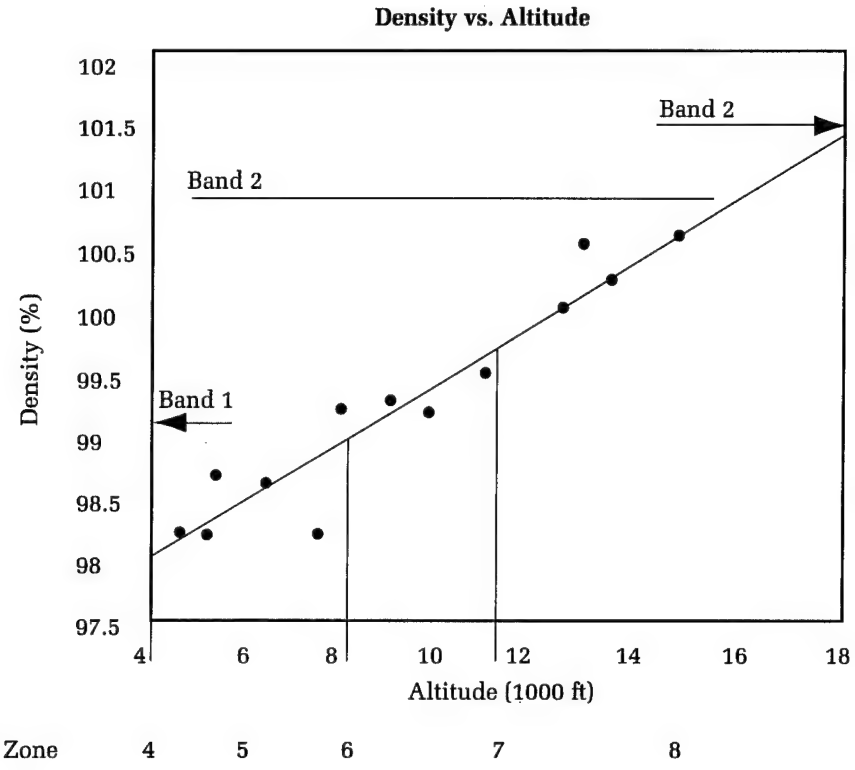


Figure 9. *Density estimation.*

altitude zones associated with the Met Message. The density value to be entered in the Met Message table for a given zone is determined by the ordinate from the linear curve corresponding to the midpoint altitude of that zone. Thus, the solid vertical lines in Figure 9 are associated with density entries in the Met Message table. Entries for wind speed and direction are determined in an entirely analogous way.

Conclusions

A concept for using projectile track data as a source of information from which to increase the accuracy of gun fire control has been presented. The concept depends critically on a procedure for processing the track data to extract the information required. One example of such a procedure is the main topic of this article. It is based on a Gauss-Newton method for obtaining a least-squares fit of a projectile flight model to the projectile track data. The procedure has been described in some detail, accompanied by a discussion of some important implementation issues.

The overall effect of the procedure is to bring projectile trajectories computed as part of gun fire control into alignment with the actual trajectories. This characteristic of the procedure allows the projectile to be placed on the aim point with greater accuracy than would otherwise be the case. This tends to minimize sensitivity to inaccuracies in the computational model of projectile dynamics used in ballistics computations. Though the projectile model can never be a totally precise representation of reality, the procedure tends to select a parameter set that makes the calculated trajectories as close to the actual ones as can be achieved with the model in use.

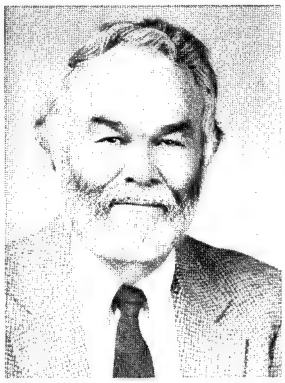
The implementation of this gun calibration procedure in the Mk 34 Gun Weapon System on the Arleigh Burke class of AEGIS destroyers has been described. Some performance results for this implementation were presented to illustrate the practical benefits.

References

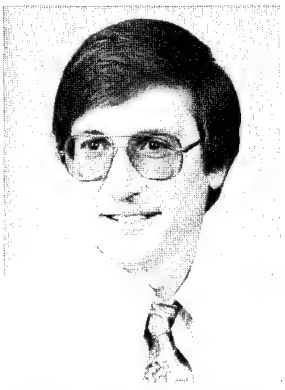
1. McShane, E. J., Kelly, J. L. and Reno, F. V., *Exterior Ballistics*, University of Denver Press, Colorado, 1953.
2. Sorenson, H. W., *Parameter Estimation*, Marcel Dekker, New York, 1980.
3. Gill, P. E., Murray, W. and Wright, M. H., *Practical Optimization*, Academic Press, New York, 1981.
4. Dennis, J. E. and Schnabel, R. B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, New Jersey, 1983.

The Authors

DONALD de FOREST BOYER received his B.A. in 1958 and M.A. in 1963, both in mathematics from George Washington University. He worked with Dr. Price on the conceptual development and the numerical implementation of the gun-calibration procedure. He has 8 years' experience in the SLBM Research and Analysis Division, applying trajectory estimation techniques to the analysis of the Trident Missile test flight data. Mr. Boyer spent 2-1/2 years in the Combat Systems Technology Branch and is currently working on weapons system simulations in the System Engineering Branch of the AEGIS Ship Combat Systems Division. Mr. Boyer is an Adjunct Professor of Mathematics at Virginia Polytechnic Institute and State University, and has recently had a paper on optimization accepted for publication in the Naval Logistics Research Quarterly.



EDWARD L. PRICE received a B.S. degree in engineering physics from the University of Tennessee at Chattanooga in 1962, an M.S. degree in electrical engineering from Virginia Polytechnic Institute and State University in 1968, and a Ph.D. degree in electrical engineering from Johns Hopkins University in 1977. He has over 20 years' experience in concept development, analysis, and design relating to weapon control and fire control for naval gun and missile systems. He participated in the development of track processing and gun-system calibration algorithms used in various models of the Mk 160 Gun Fire-Control System employed by several U.S. Naval gun systems, including the Mk 34 Gun Weapon System that is part of the combat suite on the Arleigh Burke class of AEGIS destroyers. Dr. Price is currently employed as a Technical Fellow by the Armament Systems Division of United Defense, L.P. He holds membership in the following academic honor societies: Phi Kappa Phi, Sigma Pi Sigma, and Eta Kappa Nu.



DOUGLAS O. HAAS received his B.S. degree in physics from Roanoke College in 1983. He spent 10 years as a member of the Software Engineering Group of the Weapon Control Systems Branch, where he was involved in a number of areas on the Mk 34 Gun Weapon System, including the gun-calibration procedure. More recently, he has worked as the NSWCDD software technical lead for the Ship Self-Defense System (SSDS) Demo project. He is currently working with the Software Support Activity (SSA) effort for SSDS. As a member of the SSDS Demo team, he received the NSWCDD Award of Merit for Group Achievement in 1993.

Robust Flight Control for Surface-Launched Tactical Missiles

John E. Bibel, D. Stephen Malyevac, and Ernest J. Ohlmeyer

Due to increasingly difficult threats, current air defense missile systems are pushed to the limits of their performance capabilities. To defend against these more stressing threats, surface-launched tactical missiles require greater maneuverability, faster response time, and increased robustness against more severe environmental conditions. One of the most critical missile system elements is the flight control system, as its time constant is typically half of the total missile system time constant. Conventional autopilot design techniques have worked well in the past, but in order to satisfy more stringent design specifications, new design methods are necessary. This article addresses robust control techniques (in particular, H-Infinity Control and Mu-Synthesis) and their application to the design of flight control systems for surface-launched tactical missiles. This article reviews the missile autopilot design problem, presents descriptions of the design methods and example autopilot designs using these methods, and compares the design approaches and results. The results show that applying robust control techniques to the design of missile autopilots greatly improves the stability and performance robustness characteristics of the flight control system, which increases the effectiveness of the overall missile system.

Introduction

The defense of U.S. Navy ships becomes increasingly difficult as the sophistication level of projected threats increases. These threats include such difficult targets as sea-skimming antiship missiles and tactical ballistic missiles, all of which may be launched from a variety of platforms. As the threats become more advanced, they typically become faster and more maneuverable, fly evasive flight profiles, employ enhanced guidance systems (such as multimode seekers), and use more stealthy airframes (giving lower radar cross section). In addition, the enemy may use countermeasures (electronic or otherwise) to further reduce the defensive weapon system's capabilities. Those threat characteristics lead to smaller radar detection ranges, reduced reaction time to engage the threat, restricted battle space, and small interceptor homing times, which limit the performance of the interceptor weapon system.

The projected threat characteristics are used in system-level studies that prescribe requirements upon the defensive missile weapon system. The missile system capabilities must increase to keep pace with the evolving threat capabilities. Technologies such as advanced seekers, improved guidance and control systems, stronger lightweight structures, more powerful propulsion, and innovative

airframe concepts must advance in order to meet these requirements. Reductions in the autopilot/airframe time constant is a critical element, since it accounts for roughly half of the overall missile system time constant and determines the speed and agility with which the defensive missile responds to the more maneuverable threat.

Conventional autopilot design practice is to decouple the airframe dynamics into separate single-input, single-output (SISO) channels (pitch, yaw, and roll) so that classical control techniques may be applied.[1-5] Gains of a specified feedback control loop structure are set, and the performance and stability characteristics of the flight control system are checked in a repeated process until the design requirements are satisfied. However, due to the more stringent performance demands, this conventional autopilot design approach may not be capable of achieving the fast response time, stability margin, and flexible-body mode attenuation requirements.

In order to achieve the faster response times and the improved robustness qualities prescribed by the system requirements, new techniques are necessary to design the missile autopilot. Recent advances in robust feedback control, in particular the H-infinity ($H\infty$) and Mu (μ)-synthesis techniques, have allowed for the design of very high performance missile and rocket flight controllers.[6-12] This article summarizes the role of the missile flight control system, describes both the classical control and μ-synthesis design approaches, and compares both techniques in the design of a typical tactical homing missile autopilot.

Role of Missile Autopilots

Before discussing modern robust control techniques and their application to autopilot design, the role of the autopilot subsystem in the missile homing loop and some basic control system functional requirements are reviewed. A diagram of the missile homing loop is shown in Figure 1. The basic elements of the homing loop are:

- a sensor to track the target
- signal processing and noise filters to condition the measurements and reduce errors
- a guidance algorithm to generate steering commands for the autopilot
- an autopilot/flight control system to convert these commands into actual missile maneuvers
- the kinematics of the missile-target engagement.

In this article, the discussion will be limited to the autopilot/flight control system.

The autopilot takes commands from the guidance computer and produces fin deflection commands that are sent to the actuators. The control surface actuators then operate on these commands, producing fin movements that generate aerodynamic forces and moments that maneuver the missile. In order to satisfy the requirements of the guidance system and improve airframe stability, feedback must be used to modify the airframe's dynamics. Accelerometers and rate gyros provide measurements of the airframe's response. These signals are fed back to the flight control system and, along with the guidance commands, contribute to new fin

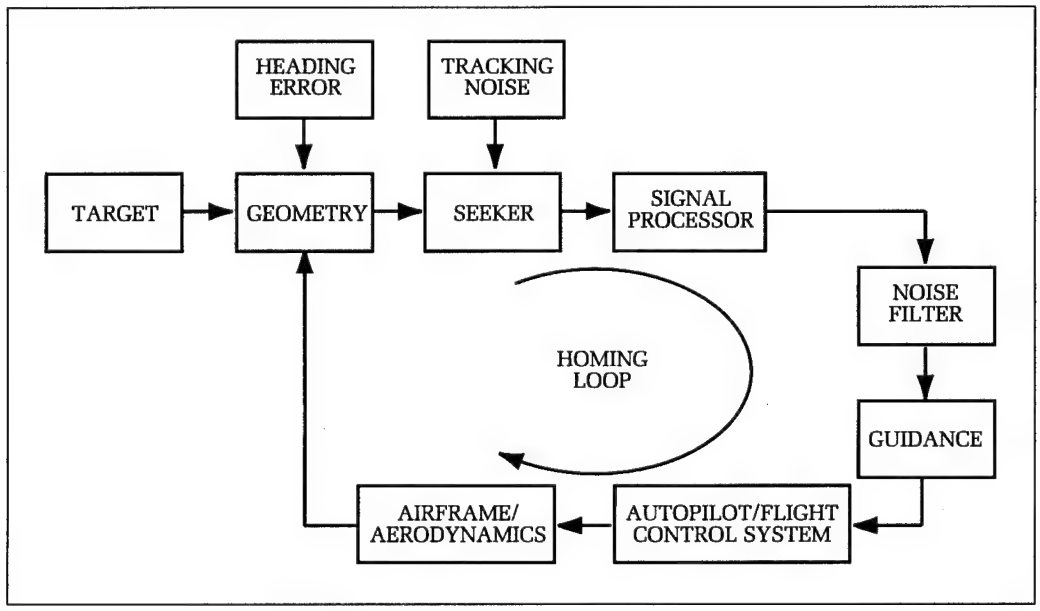


Figure 1. *Missile homing loop.*

Glossary of Symbols and Definitions

K_A, K_I, K_R, K_0 – accelerometer gain, integrator gain, rate loop gain, DC gain of autopilot

M – system transfer matrix

Δ – model uncertainty (real rational, stable transfer matrix)

ω – frequency

$\lambda(M)$ – eigenvalues of M

$\sigma(M)$ – singular values of M

$\bar{\sigma}(M)$ – maximum singular value of M

$\underline{\sigma}(M)$ – minimum singular value of M

$\| x \|_n$ – n-norm on the vector (or square matrix), x

H_2/H_∞ – control theories/methodologies, based on the use of the $2/\infty$-norm defined on the Hardy space of rational and stable transfer functions

$\mathbb{R}$, $\mathbb{C}$ – real, complex

: – "such that" or "where"

$\in$ – "is a member of" or "is contained in"

$\forall$ – "for all"

diag[A] – diagonal matrix,

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Example: $\Delta_{aug} = \{diag(\Delta, \Delta_{n+1}) : \Delta \in \mathbb{\Delta}\}$
reads that the augmented Δ block Δ_{aug} is equal to the diagonal matrix $\begin{bmatrix} \Delta & 0 \\ 0 & \Delta_{n+1} \end{bmatrix}$ where Δ is contained in the set $\mathbb{\Delta}$.

Supremum (sup) Consider a function, *f*, that is bounded from above by (possibly) several upper bounds b_{min}, ..., b_{max}. If the function, *f*, has a definite maximum, then the supremum of *f* is the maximum of *f* (*sup f = max f*). Otherwise, the supremum of *f* is the "least upper bound" b_{min} *of f*. In an analogous way, the infimum (inf) represents the maximum lower bound of *f*.

Spectral Radius Magnitude of the maximum eigenvalue of the matrix, **M**, $max\ |\lambda_i(M)|$.

deflection commands. The autopilot must track the guidance commands and control the missile to minimize the miss distance between the missile and the target.

While the specific autopilot requirements are a function of the airframe and the particular flight conditions, there are some general requirements that all autopilots must meet. Stability of the autopilot/airframe must be maintained, and the autopilot must have sufficient speed of response. The speed of response is the time it takes the autopilot/airframe to reach its commanded maneuver. Adequate damping is required for good response. An underdamped autopilot causes a large overshoot and an oscillatory response, while an overdamped autopilot gives sluggish response. The autopilot must provide tolerance (robustness) to unknown variations in the design model (the plant) and disturbances. In classical control, open-loop phase and gain margins are conventionally used as a measure of this tolerance. Each gain and phase margin represents the amount of arbitrary gain and phase lag that can be tolerated at a given loop breaking point. The larger the gain and phase margin, the more robust the system is to uncertainty and change. Another important consideration in the autopilot design is the frequency response of the autopilot/airframe. The autopilot bandwidth must be fast enough to respond to the guidance commands, while remaining out of the frequency ranges associated with the noise content in the guidance command, the missile flexible-body dynamics, and the other components of the missile flight control system; e.g., sensors and actuators. Also, the flight control system must meet these requirements over the entire flight envelope. The next section will discuss how conventional autopilots are designed to meet these requirements.

Conventional Autopilot Design

The conventional approach to designing missile autopilots is to use classical control theory techniques to compensate the airframe dynamics. Several control loop structures or topologies exist for various practical applications. One accepted autopilot topology is the three-loop autopilot shown in Figure 2.

The three-loop autopilot provides a fast time response for both stable and unstable airframes. Using a three-loop autopilot allows specified values of time constant, damping ratio, and crossover frequency to be obtained simultaneously. Details on the design approach discussed in this section can be found in References 1 through 5.

As shown in Figure 2, the three-loop autopilot uses acceleration and pitch rate measurements in three feedback loops:

- the accelerometer feedback loop
- the rate gyro feedback loop
- a synthetic stability loop.

The synthetic stability loop is a feedback loop using the integral of the rate gyro signal. The integral of the pitch rate is proportional to the angle of attack over short time intervals. This feedback signal tends to stabilize perturbations in angle of attack.

The autopilot consists of electronic gains and an integrator. The influence of the gains in the autopilot can be approximately summarized as follows:

- the rate loop gain, K_R, determines the crossover frequency
- the integrator gain, K_I, sets the damping
- the accelerometer gain, K_A, adjusts the time constant
- the forward path gain, K_0, provides unity DC transmission

This allows the designer to independently set the best time constant for a desired crossover frequency and damping ratio. As

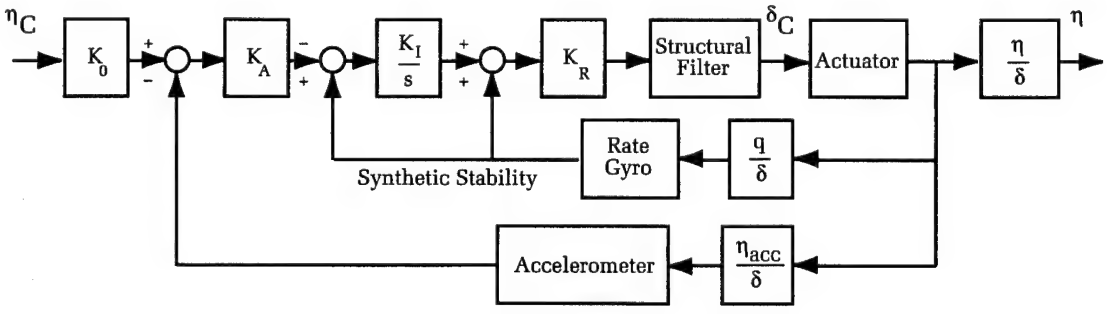


Figure 2. *Conventional three-loop pitch autopilot.*

the definition of the autopilot control law is already set, the major task of the autopilot design is to determine the autopilot gains K_A, K_I, and K_R in such a way as to optimize performance and robustness. Classical linear control theory is usually used for this purpose.

In order to use classical linear time-invariant control techniques to design the autopilot, the nonlinear aerodynamics must first be linearized. The linearization is done at a trim reference flight condition. When the missile is at trim, the sum of all the moments about the center of gravity of the missile is zero. There is a certain deflection of the control surfaces and angle of attack that will maintain the trim condition at a particular flight point. Using the linearized aerodynamics, the equations of motion for small perturbations about this trim reference condition are written. These linear perturbational dynamics are used to design the autopilot since the missile flies a large majority of its flight in a trim condition.

It is assumed that each of the autopilot channels can be separated into SISO control channels, and that the gains for each channel can be independently derived. Additional cross coupling compensation must then be added to decouple each channel. The derivation of this coupling compensation is beyond the scope of this article, and only the design of the pitch channel autopilot is discussed. The instrumentation and actuators are assumed to be perfect initially. This is a valid assumption for preliminary design since the dynamics of these elements are considerably faster than that of the airframe. Once the gains are derived, these effects are included, and the stability and performance requirements are checked. Initially, it is assumed that the airframe is a rigid body and that the flexible modes can be neglected. Later, notch filters can be added to account for the flexible-body dynamics.

The transfer function from commanded to achieved acceleration is written in terms of the perturbational dynamics and the autopilot gains. An ideal set of dynamics is derived in terms of the desired crossover frequency, damping ratio, and time constant. Equating the coefficients of the resulting third-order characteristic equation with the ideal dynamics yields three equations in the three unknown autopilot gains. The three autopilot gains (K_A, K_I, and K_R) are found by solving this system of equations. The DC gain, K_0, is then set to give unity transmission (i.e., zero steady-state tracking error).

After the gains are calculated, classical performance and stability analyses (such as step and frequency response methods) are performed on the airframe/autopilot system to verify that the requirements are met. As mentioned above, notch filters are added to the forward path of the autopilot to cancel the first and, if necessary, the second bending modes. Using a notch filter to attenuate the bending mode dynamics is called gain stabilization. Other methods that can be used to eliminate structural mode effects include:

(1) reduction of the autopilot gains; in particular, K_R, which lowers the open-loop crossover frequency and, consequently, decreases the autopilot bandwidth (undesirable from the standpoint of performance);
(2) proper choice of the gyro location, hinge line location, or the control surface center of gravity; and
(3) phase stabilization.

The autopilot is required to function over the entire flight envelope; however, the aerodynamics change appreciably with flight condition (i.e., altitude, missile speed, angle of attack, etc.). Therefore, the aerodynamics must be linearized and the autopilot gains recalculated for a wide range of flight conditions that span the entire flight envelope. The resulting set of gains are then scheduled as a function of flight condition. Onboard instruments are required to determine the current flight condition.

Robust Control

Robust control refers to an approach to control system design that builds in tolerance to differences between the real physical system to be controlled and the mathematical model used to design the controller. Mu-synthesis is one approach to robust control that minimizes the effects of specified variations on the nominal mathematical design model, while satisfying specified performance requirements. Before describing the Mu-synthesis approach, the concepts of robustness and uncertainties are highlighted, and various modern control techniques are identified.

Robustness

Feedback controllers are necessary in the design of most dynamic physical systems. The control system ensures stability of the closed-loop system, rejects the influence of

disturbances, and allows the system to achieve certain performance objectives. The controller must accomplish these goals in the presence of modeling errors and uncertainties.

The design of feedback controllers is based on some nominal model of the physical system (or plant). However, it is impossible to perfectly model the dynamics of most physical systems mathematically. Reasons for modeling errors and plant uncertainties include:

(1) partial or imperfect knowledge of the system;
(2) simplifying assumptions in the mathematical description, such as neglected high frequency dynamics;
(3) linearization of a nonlinear dynamical model; and
(4) time-varying plant behavior.

As a result, the model is an approximation and will not exactly match the true system's behavior. If the controller is designed to work with the nominal model and does not account for the possible deviations of the true system from the nominal model, the system may become unstable or not be able to meet the performance goals. Tolerance to such model uncertainties and disturbances is termed robustness.

Uncertainties

In general, there are two main classes of uncertainties: structured and unstructured. Structured uncertainties affect the plant in some known way. There is no interaction or coupling among structured uncertainty elements. Structured uncertainties can arise as parameter variations due to inaccurate knowledge of the coefficients in the differential equations that describe the physical system.

For example, consider a system modeled as a second-order differential equation

$$\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = u \qquad (1)$$

If the coefficients a and b of Equation (1) are subject to variations δ_a and δ_b relative to their nominal values, then Equation (1) becomes

$$\frac{d^2x}{dt^2} + (a+\delta_a)\frac{dx}{dt} + (b+\delta_b)x = u \qquad (2)$$

The nominal and perturbed systems are illustrated in block diagram form in Figures 3(a) and 3(b), respectively. The uncertainties δ_a and δ_b can be represented in the block diagonal (i.e., structured) form, as depicted in Figure 3(c).

Unstructured uncertainties have no particular structure in the plant model and are used to represent unspecified dynamics; they also allow for interactions (or coupling) among the assumed uncertainty elements. Unstructured uncertainties arise from many sources, such as unmodeled high frequency dynamics, nonlinearities, simplifying assumptions, and other neglected phenomena. Because uncertainties are characterized mathematically with expected bounds, unstructured uncertainties can be represented as the set of transfer functions

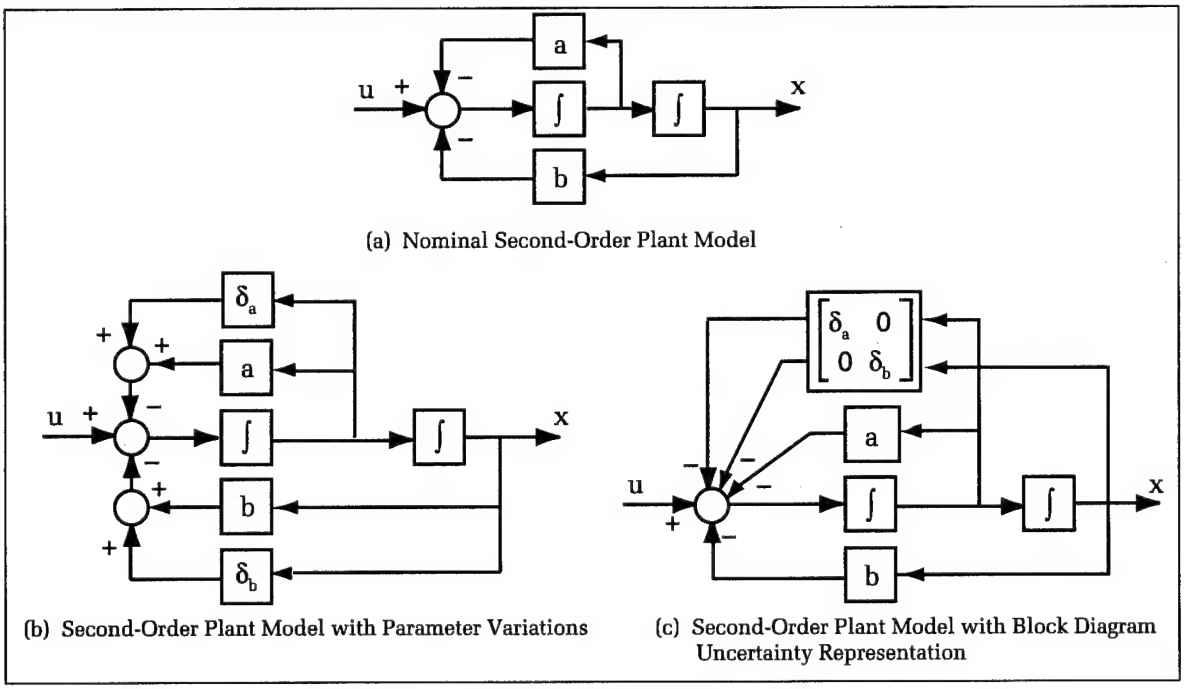
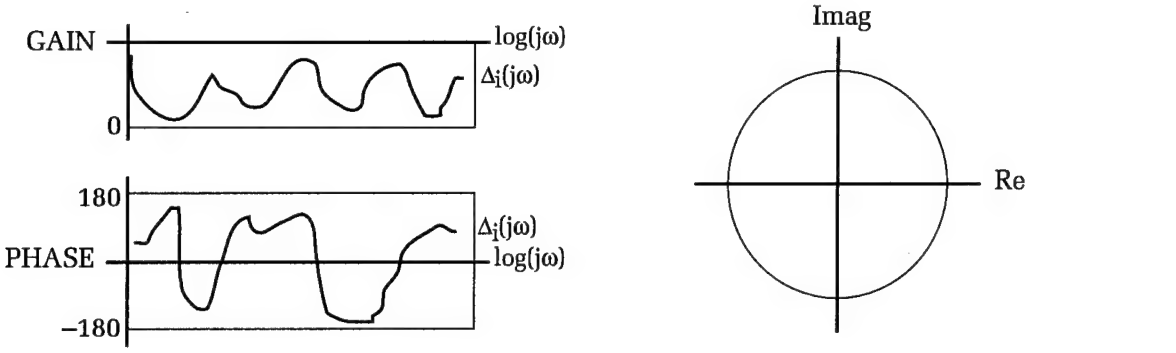


Figure 3. *Example of structured uncertainty.*



(a) Bode Representation of Unstructured Uncertainty (b) Nyquist Representation of Unstructured Uncertainty

Figure 4. *Unstructured uncertainty descriptions.*

contained in a disk on the imaginary plane (Nyquist perspective), or as those with bounded gain and unknown phase characteristics (Bode viewpoint). This is conceptualized in Figure 4.

In almost all applications, the uncertainty model of the physical system will be a mixture of structured and unstructured uncertainties. This is known as mixed model uncertainty. In general, the uncertainty elements may be real or complex, scalars or matrices, fixed or functions of frequency.

Modern Robust Control Techniques

Classical control design practice accounts for model uncertainty by requiring adequate gain and phase margins based on Bode and Nyquist analysis. Such an approach treats uncertainties only in a general sense without specifying the form of individual uncertainty models. As a result, an accurate measure of robustness is not produced.

Since the 1970s, many modern robust control methods have evolved to handle the effects of model uncertainties in the design of controllers, as discussed in References 13 and 14. These include such methods as:

- Linear Quadratic Gaussian/Loop-Transfer Recovery (LQG/LTR) approach,
- loop-shaping techniques,
- H_∞ and μ methods,
- game theoretic or minimax methods,
- Lyapunov robust stabilization,
- the Quantitative Feedback Theory of Horowitz, and
- robust eigenstructure assignment approaches,

to name a few. Mu-synthesis is a multiple-input, multiple-output (MIMO) robust control design technique that combines H_∞-control with μ-analysis in an iterative, two-step optimization process. In the next section, the H_∞ and μ methods of robust control[15-25] are reviewed.

H_∞ and μ Methods of Robust Control

Design and Analysis Framework for Robust Control

For the application of the H_∞ and μ methods, the control design and analysis problem is cast into the following framework. The physical system is modeled as a finite-dimensional, linear, time-invariant system. This model includes a description of the uncertainties and their structure, and is illustrated in Figure 5. The uncertainties are represented by Δ, **P** is the generalized plant, and **K** denotes the controller. Any linear combination of inputs, outputs, perturbations, and controller can be arranged to fit this interconnection framework. The input vector, **d**, denotes all external disturbances acting on the plant, such as commands, physical disturbances and sensor noise. The output vector, **e**, contains, in particular, the regulated performance variables. Feedback measurements are denoted by the vector, **y**, and **u** represents the control vector that is input to the plant actutors. The input disturbances, output errors,

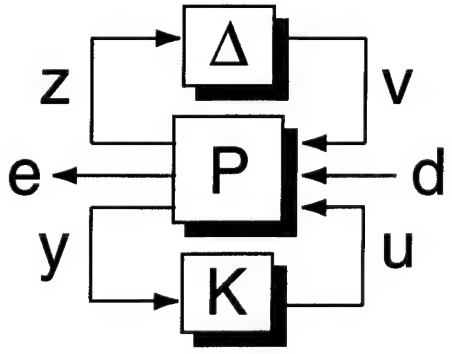


Figure 5. *Framework for robust control.*

Brief Description of Norms

Norms basically are a measure of the "size" of some element. These elements could be vectors, matrices, scalar or vector signals, and even transfer functions of a system. Also, several norms can be defined for each of these elements, such as the 2-norm or the ∞-norm. Specifically, the 2-norm and the ∞-norm are of relevance here.

Consider a scalar function, $f(t)$, of some real variable, t. This function could represent some time signal. In general, we can define a p-norm of that signal, $\|f(t)\|_p$, as

$$\|f(t)\|_p = \left\{ \int_{-\infty}^{\infty} |f(t)|^p \, dt \right\}^{1/p}$$

According to this definition, then the 2-norm of $f(t)$, $\|f(t)\|_2$, is

$$\|f(t)\|_2 = \left\{ \int_{-\infty}^{\infty} |f(t)|^2 \, dt \right\}^{1/2}$$

Thus, the 2-norm of $f(t)$ describes square-integrable functions of time and characterizes the root-mean-square values of the signal. The square of the signal 2-norm is

$$\|f(t)\|_2^2 = \int_{-\infty}^{\infty} |f(t)|^2 \, dt$$

and is defined as the "energy" contained in $f(t)$. Thus, the 2-norm of a signal can also be interpreted as a function (in particular, the square root) of the amount of energy in $f(t)$.

The ∞-norm of $f(t)$, $\|f(t)\|_\infty$, turns out to be the maximum of $f(t)$ over t, or

$$\|f(t)\|_\infty = ess \sup_t f(t)$$

where ess sup denotes the "essential supremum," which basically means the maximum value of the signal. This norm specifies the maximum (or peak) magnitude of the signal, $f(t)$.

If we let $x(t)$ represent a vector of signals that are functions of t, the 2-norm of $x(t)$ is defined as

$$\|x(t)\|_2 = \left\{ \int_{-\infty}^{\infty} trace\left[x(t)x^T(t) \right] dt \right\}^{1/2} = \left\{ \int_{-\infty}^{\infty} x^T(t)x(t)dt \right\}^{1/2}$$

Thus, the 2-norm of $x(t)$ is found by integrating the sum of the squared elements of the time response and can be interpreted as the square root of the energy of the vector-valued signal, $x(t)$.

Next, we let $G(s)$ represent the transfer function (matrix) of some system G. The 2-norm of $G(s)$ can be found by integrating the sum of the squared elements' frequency response magnitudes,

$$\|G(j\omega)\|_2 = \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} trace\left[G^*(j\omega)G(j\omega) \right] d\omega \right\}^{1/2}$$

where $G^*(j\omega)$ is the complex conjugate transpose of the frequency response matrix. The ∞-norm of $G(s)$ is essentially the peak maximum singular value of $G(j\omega)$ over frequency,

$$\|G(j\omega)\|_\infty = ess \sup_\omega \bar{\sigma}[G(j\omega)]$$

Thus, $\|G(j\omega)\|_\infty$ can be interpreted as the maximum "gain" of the transfer function evaluated along the imaginary axis.

and uncertainties are normalized to unity bounds ($\|\mathbf{d}\|_2 \le 1$, $\|\mathbf{e}\|_2 \le 1$, and $\bar{\sigma}(\Delta) < 1$) by frequency dependent weights, which are incorporated into the generalized plant model, **P**.

Robust Compensator Synthesis and H∞ Control

For robust controller synthesis, the design problem is to find a controller that stabilizes the closed-loop system and provides tolerance to model uncertainties. The control problem's performance objective can be stated in terms of norms. For example, the application of the 2-norm to the control problem leads to H_2 Control, which is commonly known as LQG control. However, LQG does not guarantee robust stability, and attempts to modify the LQG approach to recover robustness (LQG/LTR) have had difficulties (in particular, with nonminimum phase systems and multiple uncertainties). On the other hand, the ∞-norm possesses attractive properties for application to robust control. First, the ∞-norm possesses the *submultiplicative property*, which is essential in establishing robustness of uncertain systems. Consider the cascade of two systems, M_1 and M_2, as given by

$$y(s) = M_2(s)M_1(s)u(s) \qquad (3)$$

The infinity norm satisfies the inequality

$$\|M_2M_1\|_\infty \le \|M_2\|_\infty \|M_1\|_\infty \qquad (4)$$

This inequality says, in particular, that the energy gain through both systems M_1 and M_2 is less than or equal to the product of the energy gain through each individual system. In contrast, the 2-norm does not satisfy the submultiplicative property of Equation (4). To establish robust stability of an uncertain system, the gain or amplification that occurs through the loop is important (using either MIMO Nyquist theory or the Small Gain Theorem). That gain can be bounded by using infinity norms but not by using 2-norms. Another useful property of the ∞-norm for control is derived from its definition. The infinity-norm of the system M, $\|M\|_\infty$, can be interpreted as the worst case, square root energy of the output for unit energy input to M. Thus, the infinity-norm is useful if the performance objective is to minimize the transfer of disturbance energy through the system. For these reasons, the ∞-norm is a powerful tool for robust control and led to the initial development of H_∞ control theory (Zames).[24]

To formalize the H_∞ control problem, consider the control framework in Figure 6. In this diagram, **P(s)** represents the generalized plant transfer matrix, and **K(s)** is a linear transfer matrix description of the controller. The plant, **P**, has two inputs (**W** and **u**) and two outputs (**Z** and **y**); and absorbs weighting functions on the (disturbance) inputs, **W**, and the (performance) outputs, **Z**. Weighting functions on the uncertainties are also absorbed into the generalized plant, **P**. Closing the feedback control loop, the transformation (or mapping) of the input, **W**, to the output, **Z**, T_{ZW}, is called the lower linear fractional transformation, $F_l(P,K)$,

$$Z = T_{ZW}W = F_l(P,K)W \qquad (5)$$

where

$$F_l(P,K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21} \qquad (6)$$

Here, P_{ij} are the four submatrices formed by partitioning the matrix P.

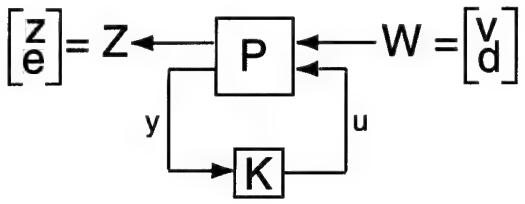


Figure 6. *Controller synthesis framework.*

The H_∞ control problem is then to find a real-rational, proper controller, K, that makes the closed-loop system internally stable and that minimizes the infinity-norm of the input/output map, T_{ZW}; i.e.,

$$\min_K \|T_{ZW}\|_\infty = \min_K \|F_l(P,K)\|_\infty \qquad (7)$$

Thus, H_∞ control yields a closed-loop system, M, whose ∞-norm is minimized over all controllers that stabilize the closed-loop system. The H_∞ controller can have several interpretations. In one view, H_∞ control minimizes the energy gain between the exogenous inputs, **W**, and the generalized error outputs, **Z**. Another interpretation is that H_∞ control minimizes the largest of the maximum singular values of the transformation, T_{ZW}, found over all frequencies. As such, the solution minimizes the maximum possible amplification of the (disturbance) inputs through the system. So, in general, H_∞ control may be thought of as a "worst-case" minimization design approach.

The H_∞ optimal control (central) solution is obtained via the solutions of two algebraic Riccati equations.[22-23] The controller, K, has the same order as the generalized plant and also has a separation structure similar to that

for LQG (H_2) controllers. That is, the H_∞ controller contains both controller gains and observer (filter) gains. Several solution methods to the H_∞ problem have been developed; however, the state-space solution given in References 22 and 23 is the most computationally practical.

Analysis and the Structured Singular Value

For analysis purposes, the control loop is closed, and the controller is incorporated in with the generalized plant to form the closed-loop system transfer matrix, **M**. The analysis framework is shown in Figure 7, where Δ represents the model uncertainty. The system, **M**, has two inputs **(v, d)** and two outputs **(z, e)**. The input/output map from **d** to **e** can be expressed as the upper linear fractional transformation, $F_u(M,\Delta)$

$$e = F_u(M,\Delta)d \qquad (8)$$

where

$$F_u(M,\Delta) = M_{22}+M_{21}\Delta(I-M_{11}\Delta)^{-1}M_{12} \qquad (9)$$

and M_{ij} are matrix partitions of M. Tests for stability and performance analysis are now presented.

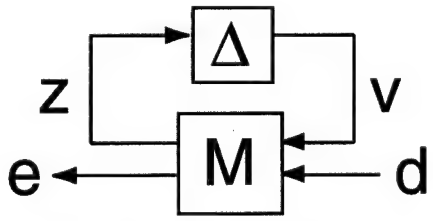


Figure 7. *Analysis diagram.*

Nominal Stability (NS). The nominal closed-loop system (i.e., no uncertainties) is said to be nominally stable if **M** is stable. This can be checked by calculating the eigenvalues of the closed-loop system, **M**. If all eigenvalues have real parts less than zero, the system is nominally stable.

Nominal Performance (NP). Let the inputs, **d**, and the outputs, **e**, be bounded: $\|d\|_2 \le 1$ and $\|e\|_2 \le 1$, and take Δ to be zero. The system satisfies nominal performance if

$$\|M_{22}\|_\infty = \sup_\omega \bar{\sigma}\left[M_{22}(j\omega)\right] \le 1 \qquad (10)$$

In essence, the system is said to satisfy nominal performance if **M** is stable and satisfies the performance requirements.

Robust Stability (RS). Robust stability refers to the stability of the perturbed closed-loop system. In general, the system is robustly stable if the closed-loop feedback system is stable in the presence of parameter variations and other model uncertainties, Δ. A discussion of the uncertainty, or Δ, block is now given.

Uncertainty Descriptions. It is simplest to consider Δ to be complex with a block diagonal structure

$$\Delta = diag\{\Delta_1,\Delta_2,\ldots,\Delta_N\} \qquad (11)$$

since this is convenient for analysis and synthesis. The individual uncertainties, Δ_i, that make up the Δ block may have one (or a combination) of several representations:

(1) scalar block – uncertainty appears only once in the real system, such as a parameter uncertainty;
(2) repeated scalar blocks – where the uncertainty appears multiple times in the real system, such as dynamic pressure; and
(3) full blocks – examples of which include multivariable neglected dynamics and the performance block.

Robust stability tests can be defined for various uncertainty descriptions, such as unstructured or structured uncertainties.

Robust Stability for Unstructured Uncertainty. In this case, Δ is a full uncertainty block. For bounded uncertainty, $\bar{\sigma}(\Delta) < 1$, the closed-loop system satisfies robust stability if

$$\|M_{11}\|_\infty = \sup_\omega \bar{\sigma}\left[M_{11}(j\omega)\right] \le 1 \qquad (12)$$

Robust Stability for Structured Uncertainty. In most practical circumstances, at least some of the structure of the uncertainty is known. If the analysis test for unstructured uncertainty given in Equation (12) were used for systems with structured uncertainty, the estimate for robust stability would be conservative. To reduce the conservatism, the structure of the uncertainty must be accounted for. In one case, assume the uncertainties are all repeated complex scalars ($\Delta = \delta_i$). The system is robustly stable if $\rho(M_{11}) \le 1$, where $\rho(M_{11})$ is the spectral radius of M_{11} (i.e., the magnitude of the maximum eigenvalue of M_{11}, $max\ |\lambda_i(M_{11})|$).

Allowing for more generality in the uncertainty structure, assume that Δ takes the form

$$\mathbb{\Delta} = \left\{ \begin{matrix} diag\left[\delta_1 I_{r_1},\ldots,\delta_i I_{r_i},\ldots,\delta_N I_{r_N},\Delta_1,\ldots,\Delta_j,\ldots,\Delta_M\right] : \\ \delta_i \in \mathbb{C}, \Delta_j \in \mathbb{C}^{m_j x m_j} \end{matrix} \right\} \qquad (13)$$

In other words, Δ takes the form of complex-valued, block diagonal perturbations. Each δ_i

scalar block is repeated r_j times, and each Δ_j full block has dimension $m_j \times m_j$. Further, the uncertainties may be members of a bounded subset of those given above

$$B\mathbb{\Delta} = \left(\Delta \in \mathbb{\Delta}: \bar{\sigma}(\Delta) \le 1\right) \tag{14}$$

Then, the structured singular value $\mu(M(j\omega))$ of a complex matrix $M(j\omega)$ with respect to the block structure Δ is defined as[19]

$$\mu\left(M(j\omega)\right) = \left\{ \begin{array}{c} \dfrac{1}{\min\limits_{\Delta \in \mathbb{\Delta}} \left[\bar{\sigma}(\Delta): det\left(I - M(j\omega)\Delta\right) = 0\right]} \\ or\ 0,\ det\left(I - M(j\omega)\Delta\right) \neq 0 \end{array} \right\} \tag{15}$$

This says that μ is the inverse of the smallest magnitude of a destabilizing perturbation of M. The structured singular value is then a measure of robustness of the model to structured perturbations. In practice, the supremum of μ over frequency gives the inverse of the minimum-sized destabilizing perturbation. Thus, the structured singular value provides a quantitative measure of how much uncertainty can be tolerated before the system becomes unstable. The system is said to be robustly stable if

$$\sup_{\omega} \mu\left(M_{11}(j\omega)\right) \le 1 \tag{16}$$

The optimization problem given in Equation (16) is not globally convex, and the computation of μ is difficult to perform. However, useful upper and lower bounds can be found for μ. For any perturbation set Δ, μ is bounded by

$$\rho\left(M(j\omega)\right) \le \mu\left(M(j\omega)\right) \le \bar{\sigma}\left(M(j\omega)\right) \tag{17}$$

This equation says that the structured singular value of $M(j\omega)$ is bounded above by the maximum singular value of $M(j\omega)$ and is bounded below by the spectral radius of $M(j\omega)$. These bounds, however, are not tight and are, therefore, not very useful for computation. To improve the bounds, we introduce the transformation matrices, U and D, where U is a diagonal, unitary $(U^*U = I)$ matrix that belongs to the set $\mathcal{U}$ $(U \in \mathcal{U})$; and the D matrix is diagonal and invertible (D^{-1} exists $\forall$ $D \in \mathcal{D}$). The sets $\mathcal{U}$ and $\mathcal{D}$ have the same structure as Δ, with the properties that μ is U-invariant and is also invariant under a similarity transformation with D; i.e.,

$$\begin{aligned} \mu(UM) &= \mu(MU) = \mu(M):\ U \in \mathcal{U} \\ \mu\left(DMD^{-1}\right) &= \mu(M):\ D \in \mathcal{D} \end{aligned} \tag{18}$$

These properties allow for a tighter set of bounds on μ,

$$\begin{aligned} &\max_{U \in \mathcal{U}} \rho\left[U(j\omega)M(j\omega)\right] \le \mu\left[M(j\omega)\right] \\ &\le \inf_{D \in \mathcal{D}} \bar{\sigma}\left[D(j\omega)M(j\omega)D^{-1}(j\omega)\right] \end{aligned} \tag{19}$$

The equality on the lower bound is true for all $M(j\omega)$ and all structures $\mathbb{\Delta}$, but the computational optimization problem of finding this lower bound has many local maxima and is, therefore, difficult to compute. The equality on the upper bound holds in some cases, but is not true in general. However, the optimization problem for the upper-bound calculation is convex, and the D-scales can be used to obtain an upper-bound value that is fairly close to the μ value in most cases. So, an approximate answer to the μ problem is

$$\begin{aligned} &\sup_{\omega} \mu\left[M(j\omega)\right] \\ &\le \sup_{\omega} \inf_{D \in \mathcal{D}} \left\| D(j\omega)M(j\omega)D^{-1}(j\omega) \right\|_{\infty} \end{aligned} \tag{20}$$

In practice, robust stability is checked using

$$\sup_{\omega} \inf_{D \in \mathcal{D}} \bar{\sigma}\left(D(\omega)M_{11}(j\omega)D^{-1}(\omega)\right) \le 1 \tag{21}$$

Robust Performance (RP). In general, robust performance means that the closed-loop system is stable and achieves the performance specifications in the presence of modeling uncertainties. The performance requirements are specified by means of weighting functions that are applied to the regulated output variables. The performance block, Δ_{n+1}, is the transfer matrix between the disturbance inputs, **d**, and the error outputs, **e**. The performance block is treated as a full uncertainty block, assuming $\bar{\sigma}(\Delta_{n+1}) \le 1$. The robust performance problem is converted into an equivalent robust stability problem by augmenting the (robust stability) uncertainty block

$$\Delta_{aug} = \left\{diag\left(\Delta, \Delta_{n+1}\right): \Delta \in \mathbb{\Delta}\right\} \tag{22}$$

The system exhibits robust performance if

$$\sup_{\omega} \mu_{\Delta_{aug}}\left(M(j\omega)\right) < 1 \tag{23}$$

provided that **d** and **e** have unity-bounded 2-norms and that $\Delta \in B\mathbb{\Delta}$. The upper bound for μ is used in practice to check for robust performance,

$$\begin{aligned} &\sup_{\omega} \mu_{\Delta_{aug}}\left[M(j\omega)\right] \\ &\le \sup_{\omega} \inf_{D \in \mathcal{D}} \bar{\sigma}\left(D(\omega)M(j\omega)D^{-1}(\omega)\right) \le 1 \end{aligned} \tag{24}$$

If the upper bound with respect to the augmented uncertainty block is less than unity over all frequencies, then the closed-loop system is guaranteed to remain stable and satisfy the performance requirements in the presence of all the specified structured uncertainties in the set Δ. Note that this upper bound test has become an ∞-norm test.

Mu-Synthesis

The ultimate goal of designing a controller is to satisfy the robust performance criterion. This goal is met by minimizing the sufficient condition for robust performance given in Equation (24) over all stabilizing controllers;

$$\inf_{K} \sup_{\omega} \inf_{D(\omega)} \bar{\sigma}\left[D(\omega)M(P,K)(j\omega)D(\omega)^{-1} \right] \quad (25)$$

This is done iteratively by solving for the controller K and the D-scales, $D(\omega)$, with the goal

$$\left\| DM(P,K)D^{-1} \right\|_{\infty} \leq 1 \quad (26)$$

The successive iterations are known as "DK Iterations."

The DK iteration process is summarized as follows. Consider some stabilizing controller, K. Closing the control loop yields the lower linear fractional transformation, $F_l(P,K) = M(P,K)$. The D-scales yield the transformation $DM(P,K)D^{-1}$. The upper bound is the infimum of the infinity-norm of this latter transformation taken over all stabilizing controllers, K, and all admissible D-scales. This is an optimization problem in the two variables K and D. For a given D, we have an H_∞ control problem. For a fixed controller K, an optimal D-scale can be computed. The D-scales are computed to minimize the upper-bound approximation. That is, the largest singular value of the transformation is minimized at each frequency as a convex optimization problem in D. The resulting $D(\omega)$ is then fit with a stable, rational transfer function with a stable inverse. This iterative process is performed until the ∞-norm of $DF_l(P,K)D^{-1}$ is close to its μ value. This iterative optimization problem is convex when either variable is fixed but is not globally convex and is not guaranteed to converge to the global optimal K and D. However, the DK approximation has been shown to be quite good in a large number of engineering applications, illustrating that the technique is very useful for the design of robust controllers.

The engineering process for control design using μ-synthesis can be described in the following steps:

(1) Formulate the model of the plant;
(2) Set up the uncertainty description for the model, including the uncertainty weights;
(3) Design the weighting functions for the inputs and performance outputs;
(4) Perform the DK iteration until convergence;
(5) Analyze the resulting design for nominal stability, nominal performance, robust stability, and robust performance; and
(6) If necessary, repeat steps 2 through 5 until the design requirements are satisfied.

The first three steps above are the most critical steps in the design process. In the formulation of the plant and uncertainty models, there is a tradeoff between how much detail from the real system to include in the design model versus the complexity of the controller that is synthesized. A controller too large in dimension may be too cumbersome to implement in a practical flight computer. Also, a design model that is overly complex in terms of its uncertainty description may result in a compensator that is extremely robust to perturbations but at the cost of performance. In addition, the selection of the input/output weights are important, because they act like "tuning knobs" in the H_∞ and μ-synthesis methods. There are many problem dependent issues regarding the design of these weights that must be considered, as discussed in Reference 25.

Missile Autopilot Design Example

The problem treated in this article is the design of a pitch plane autopilot with flexible-mode compensation using both a conventional autopilot design approach and a gain-stabilized, μ-synthesis approach. The missile airframe considered in this study is shown in Figure 8. The missile is tail controlled, with dorsals added to increase lift and maneuverability, and has a high L/D ratio. Long slender missiles have first bending-mode frequencies that limit the achievable responsiveness of the control system, unless special design procedures are followed. Figure 9 illustrates the elastic bending modes associated with the candidate short range, air defense missile analyzed in the article.

The primary objective for this autopilot design is to obtain the fastest time response possible while maintaining specified stability robustness characteristics. In particular, the

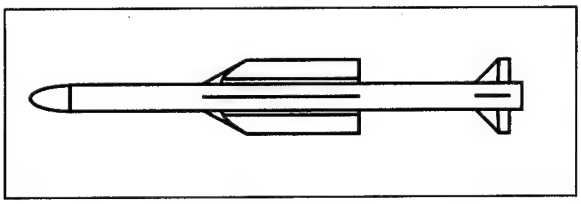


Figure 8. *Missile configuration.*

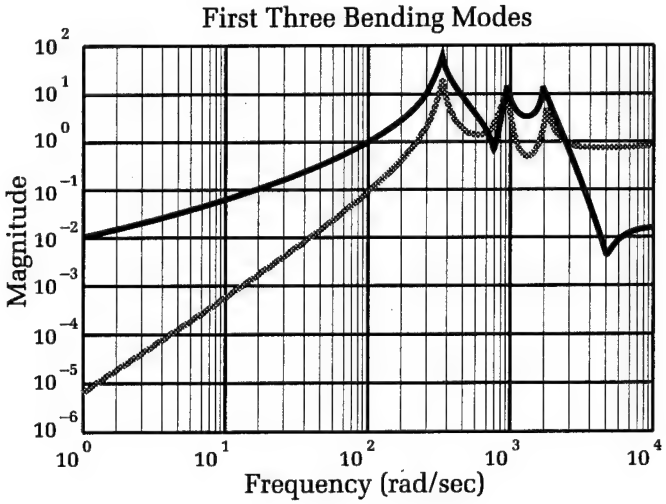


Figure 9. *Elastic-body frequency response.*

normal acceleration response attributable to a unit step command should be fast (time constant ≤ 0.15 second), while exhibiting small steady-state error (≤ 1%) and small overshoot (≤ 4%). (The time response characteristics would have been generated from systems design requirement analysis, considering the interceptor performance that must be achieved to intercept specific threats.) For stability robustness, 6 dB gain margins and 30 degree phase margins are required for the control loop opened at the actuator input, the gyro output, and the accelerometer output locations. In addition, to ensure further robustness, it is required that the Nyquist plot does not enter a ball of 0.5 radius about the critical point for the μ-synthesis design. (A 0.5 radius keep-out zone guarantees a minimum 6 dB gain margin and 30 deg phase margin from the critical point.) To achieve flexible-body compensation for a gain-stabilized design, a minimum of 20 dB of attenuation at the flexible-mode peak frequencies is required to prevent noise, saturation, or aeroelastic vibration coupling problems.

The first autopilot design considered in this article used conventional design practice. The approach used for flexible-mode compensation was the addition of a notch filter tuned to the first bending-mode frequency. This structural notch filter is added to the flight control loop as shown in Figure 2. To set the autopilot gains, the time constant of the desired characteristic equation was set to 0.08 second, the inner rate loop crossover frequency was set to 35 rad/second, and the damping ratio was set to 0.7, which allowed all the design specifications to be satisfied.

The time response to a unit step command is shown in Figure 10. For this autopilot design, the achieved time constant is 0.143 seconds. Figure 11 gives the Bode plot for the loop opened at the actuator input. This plot, as well as similar ones for the loop opened at the gyro and at the accelerometer, indicates adequate stability margins and elastic body attenuation.

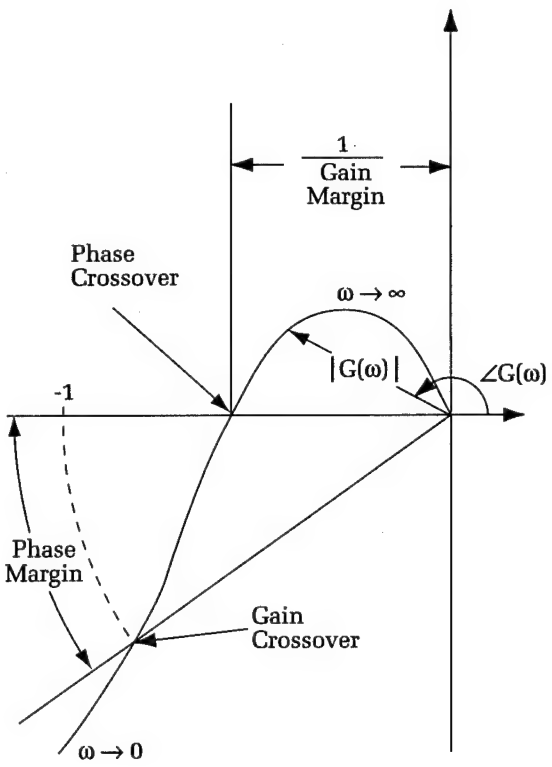
Next, μ-synthesis was applied to design the autopilot. The μ-synthesis design model is shown in Figure 12 and is built around the same linearized airframe dynamics used for the conventional autopilot design. Disturbance inputs to the design model include the acceleration command signal, rate gyro drift, and accelerometer bias. A gyro drift value of 5 deg/hr and an accelerometer bias value of 0.0001 g are used to weight those respective inputs.

The regulated output performance variables considered here are the error between the acceleration command and the achieved normal acceleration, the fin deflection angle, the actuator fin rate, and the body pitch rate. All of these outputs are normalized using frequency dependent weighting functions, which are prescribed by the performance specifications. The weight on the acceleration error, which is the sensitivity function, is generally selected as a low-pass transfer function, with its low frequency gain set equal to the inverse of the desired steady-state tracking error, its crossover frequency set approximately equal to the inverse of the desired closed-loop time constant, and its high frequency magnitude selected to limit overshoot. The sensitivity function weight was chosen as

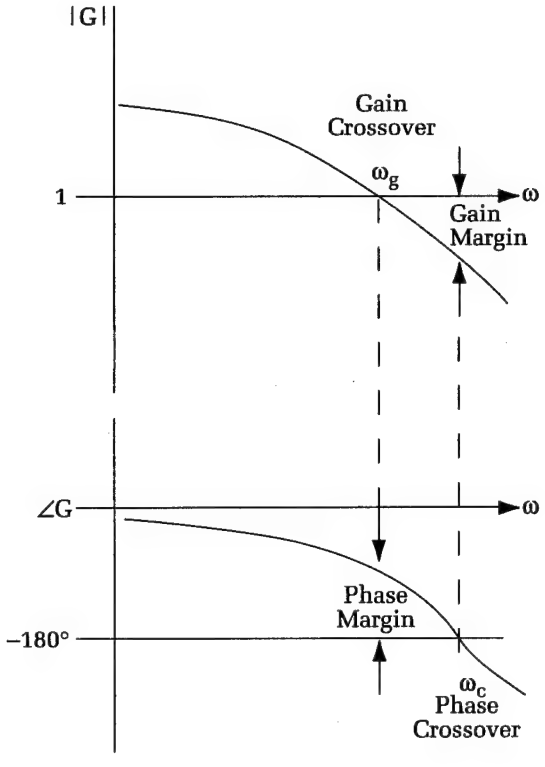
Summary Description of Bode and Nyquist Plots for Stability Analysis

Bode and Nyquist plots are classical control techniques used to analyze system stability and robustness in the frequency domain. These methods look at a system's steady-state response to sinusoidal inputs of varying frequencies. The linear, time-invariant system's response to a steady sinusoidal input is also sinusoidal, having the same frequency as the input, but displaced through a phase shift with an amplitude magnification or attenuation. This amplitude magnification or attenuation is expressed in terms of the ratio between the output and input amplitudes, and is called the amplitude ratio. The frequency response for a system is completely specified if the amplitude ratio and phase shift are known for all frequencies. Typically, a system's frequency response is represented either by two plots, one of the amplitude ratio and one of the phase angle both versus input frequency, or by a single polar plot. The two-plot representation is called a Bode plot after H. W. Bode, and the polar plot representation is named for H. Nyquist.

The figures below show typical open-loop Bode and Nyquist plots. Also indicated on these plots are the gain and phase margins. These margins represent the amount of arbitrary gain increase and phase lag that can be tolerated at the loop breaking point before closed-loop instability results. The larger the gain and phase margin, the more robust the system is to uncertainty and change. The gain margin is defined at the frequency of the phase crossover point. On the Nyquist plot, this is where the locus intersects the negative real axis and, on the Bode plot, where the phase trace crosses –180 degrees. Similarly, the phase margin is defined at the gain crossover frequency. The gain crossover frequency is where the amplitude ratio is one, as shown below.



Nyquist Plot



Bode Plot

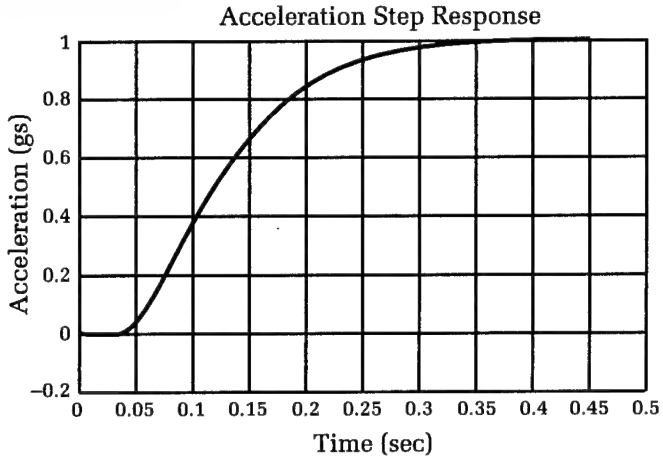


Figure 10. *Acceleration: conventional design.*

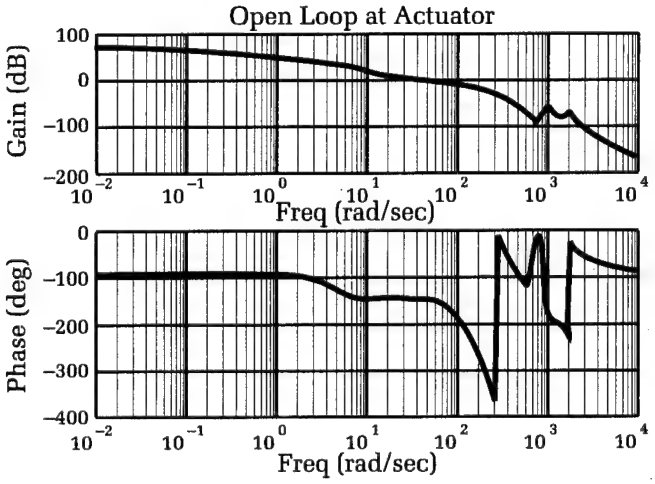


Figure 11. *Bode plot: conventional design.*

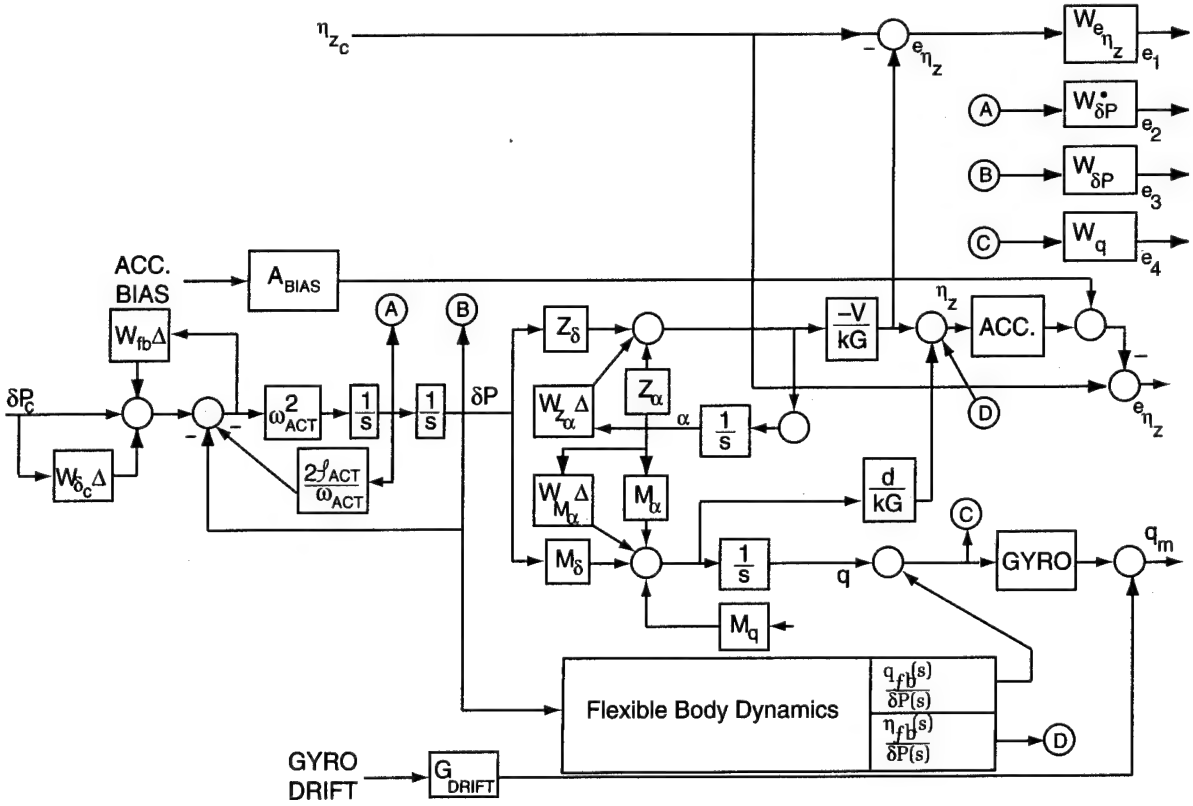


Figure 12. *Mu-synthesis design model.*

$$W_s(s) = \frac{100(0.0258s + 1)}{10.327s + 1} \qquad (27)$$

The weights on the fin angle position and the actuator fin rate were selected as constants over frequency, equal to the inverses of their respective saturation limits. This prevents these variables from saturating during flight. The body pitch rate was weighted to help achieve high frequency, control loop gain roll-off and also to help meet the high frequency attenuation goals. In contrast to the sensitivity function weight, the pitch rate weight, $W_q(s)$, is high pass, with a small gain at low frequency and a large gain at higher frequencies to suppress the effects of high frequency unmodeled dynamics and sensor noise. A second-order transfer function was used for this weight, with its low frequency gain set equal to the inverse of the assumed maximum rate of the gyro (400 deg/second), and its crossover frequency and high frequency gain tuned to satisfy the design specifications. This output performance weight is given by

$$W_q(s) = \frac{(1/400)(0.025s + 1)^2}{\left(1.78e^{-4}s + 1\right)^2} \qquad (28)$$

Four model perturbations, or uncertainties, are also considered in the design model. These include two aerodynamic coefficient uncertainties and two actuator perturbations. One of the desired controller attributes is to remain insensitive to angle-of-attack changes during flight. However, the pitching moment and normal force aerodynamic characteristics are nonlinear over the angle-of-attack range of the airframe. To account for this in the design model, multiplicative parameter uncertainties are introduced on the pitching moment and normal force aerodynamic stability derivatives with respect to angle of attack. The weights on these uncertainties reflect the percentage variation of those derivatives about their respective average values over the angle-of-attack range (M_α variation = 270% and N_α variation = 15%). To satisfy the specified high frequency attenuation and stability margin/robustness goals, feedforward (multiplicative) and feedback (divisive) uncertainties are added at the actuator input signal. A constant value weighting of 0.2 was found to be adequate for the actuator feedback uncertainty because of contributions from other weights. (A constant value at 0.5 over frequency for this weight would have strictly enforced the Nyquist 0.5 radius robustness requirement, but at the cost of performance.) For the actuator feed-forward uncertainty weight, a second-order, high-pass weight was used to help tune the controller. This uncertainty weight is given by

$$W_{unc_{ff}}(s) = \frac{(1/100)(0.049s + 1)^2}{\left(9.805e^{-4}s + 1\right)^2} \qquad (29)$$

After these weighting function and uncertainty elements are incorporated into the μ-synthesis design model, the system is arranged into the general interconnection structure diagram shown in Figure 13.

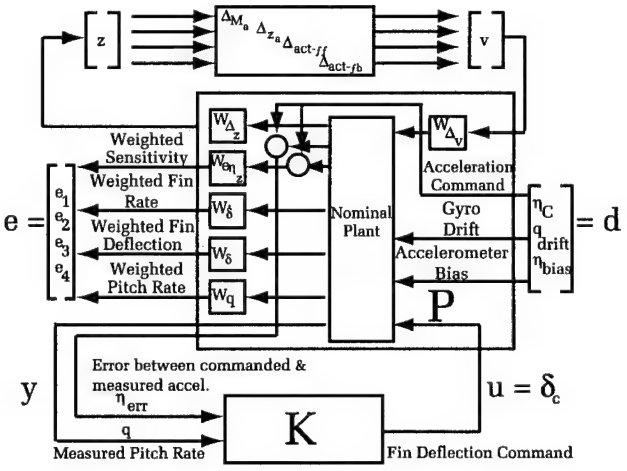


Figure 13. *Interconnection structure.*

The resulting μ-controller, obtained after two DK iterations, stabilizes the closed-loop system and is of 27th order. These states come from the generalized plant, which include the performance and uncertainty weights, and the D-scale weights. For example, four poles of the compensator are the same as those of the performance and uncertainty weights. The μ-synthesis controller builds in notches at the three body bending mode frequencies in the pitch rate channel. Note that this differs from the conventional approach, where the notch filters are added at the actuator input. The acceleration error control loop gain does not have any notches but rolls off to obtain high frequency attenuation.

The time response of the closed-loop system to a unit step acceleration command exhibits fast performance. The achieved acceleration is shown in Figure 14, where the time constant is 0.086 second, and the overshoot is approximately 2 percent. Use of the robust design approach has permitted very agile missile response despite the low frequency elastic-body bending modes. This fast response demonstrates the speed advantage of

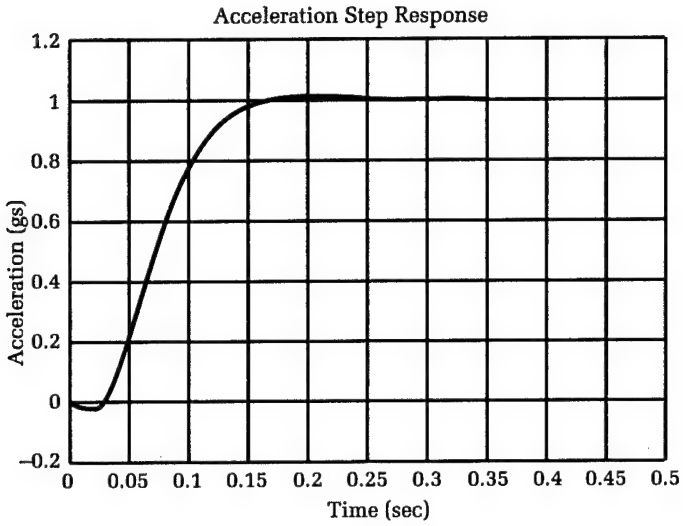


Figure 14. *Acceleration: μ-synthesis design.*

this controller over that of the conventionally designed autopilot. The flexible-body dynamics were also found not to disturb the angle-of-attack and pitch rate dynamics of the missile, as verified by simulation.

Frequency response analysis was performed at the plant input and measurement output locations. In particular, the gain margin, phase margin, body bending mode attenuation and Nyquist loop transfer requirements were checked. The Bode and Nyquist plots for the loop opened at the actuator input are shown in Figures 15 and 16, respectively. The Nyquist plot of Figure 16 is a close-up view about the critical point and shows that the loop transfer remains outside the required 0.5 disk illustrated on the figure about the critical point. These results demonstrate that the gain and phase margins are adequate, as are the high frequency attenuation goals. The frequency response requirements were also satisfied for the loop opened at the measurement output locations.

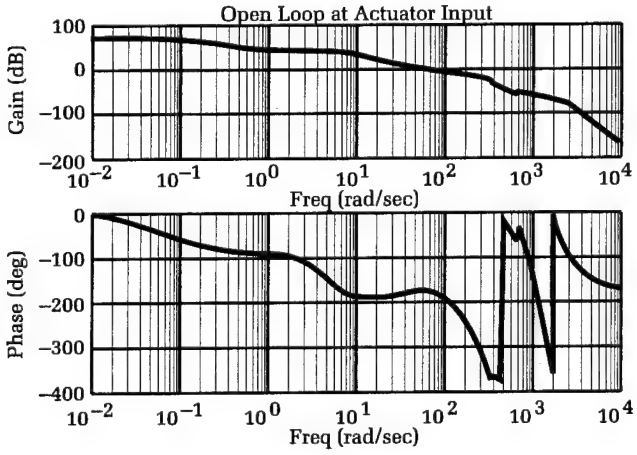


Figure 15. *Bode plot: μ-synthesis design.*

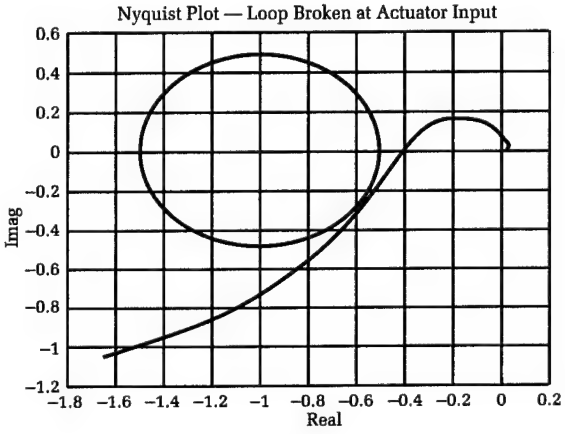


Figure 16. *Nyquist plot: μ-synthesis design.*

In addition, μ-analysis is performed to check robust stability and robust performance characteristics. Figure 17 displays the plots for nominal performance, robust stability, and robust performance, as discussed previously. The closed-loop system exhibits robust stability and performance in the presence of the uncertainties assumed in the design model (aero variations and actuator uncertainties). Thus, all the design objectives were achieved.

The 27-state controller described above is rather large to be implemented onboard a current missile's guidance and control computer. Therefore, it is desired to reduce the order of the controller as low as possible to make implementation easier. At the same time, the previously specified requirements must still be satisfied. The only exception to this is the loop transfer robustness specification on the Nyquist plot. This requirement was included in the full-order, μ-controller design to enforce good gain and phase margins on its reduced-order variant.

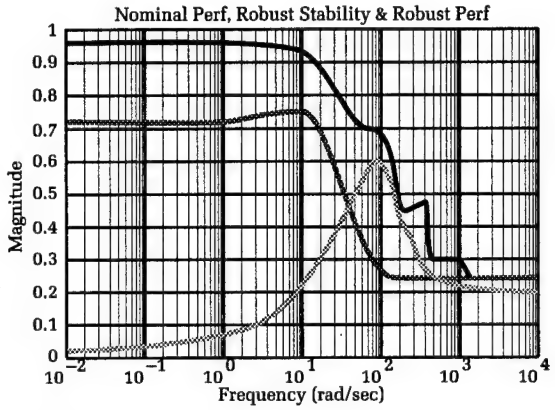


Figure 17. *Nominal performance, robust stability, and robust performance: μ-synthesis design.*

Techniques for reducing the size and complexity of μ-synthesis controllers are now readily available, and substantial reductions are usually possible with little or no sacrifice in robust performance. The controller order was reduced in the present case by inspecting the controller poles and zeros. This entailed neglecting the controller's high frequency dynamics and canceling stable pole/zero combinations located close to each other. These simplifications resulted in two decoupled third-order controllers for the autopilot acceleration error and pitch rate channels. The resulting controller transfer functions are

$$K_{\eta_{error}}(s) = \frac{-5.3}{s+0.1}\left[\frac{s+5.5}{s+6.4}\right]\left[\frac{s+80}{s+200}\right] \quad (30)$$

$$K_q(s) = 0.2\left[\frac{s+28}{s+6.4}\right]\left[\frac{s^2+2\zeta_N\omega_N+\omega_N{}^2}{s^2+2\zeta_D\omega_D+\omega_D{}^2}\right] \quad (31)$$

where $\zeta_N = 0.01$, $\omega_N = 378$ rad/second, $\zeta_D = 0.06$ and $\omega_D = 754$ rad/second. The controller in the acceleration error channel has a pole near zero to act as an integrator for zero steady-state error, a lead-lag element near crossover, and a lead-lag before the first flexible body mode. The controller for the pitch rate channel includes a lead-lag combination near the crossover frequency and a notch filter element for the first flexible body mode. Note that only the first notch filter was retained, as compared to the three notch filters for the full-order controller, since the attenuation due to roll off at the higher frequencies was adequate enough to satisfy the design requirements. Also, the two channels share only one common pole.

Analyses of the controller and of the open- and closed-loop systems verify that time response, stability margin, and high frequency attenuation requirements were satisfied. Figure 18 shows the time response of the reduced-order system to a unit step acceleration command. The response characteristics are similar to those of the full-order controller, except that an even slightly faster time constant of 0.078 second was achieved. In addition, the frequency response specifications, such as stability margins and high frequency attenuation, were satisfied, as verified by Bode plots for different loop breaking points. The robust stability and robust performance curves are shown in Figure 19. A comparison of this figure with Figure 17 shows that although the reduced-order design has some peaking in the robust stability and robust performance curves near the second body bending frequency, all the curves for the reduced-order design fall below one. This indicates that the reduced-order design satisfies all performance and robustness objectives for the assumed design model, with a substantial decrease in complexity relative to the full-order model.

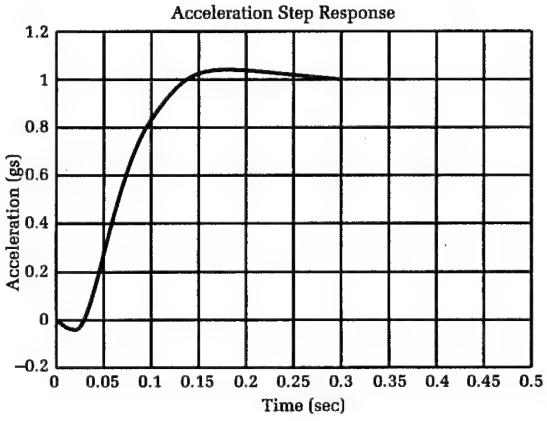


Figure 18. *Acceleration: reduced μ-design.*

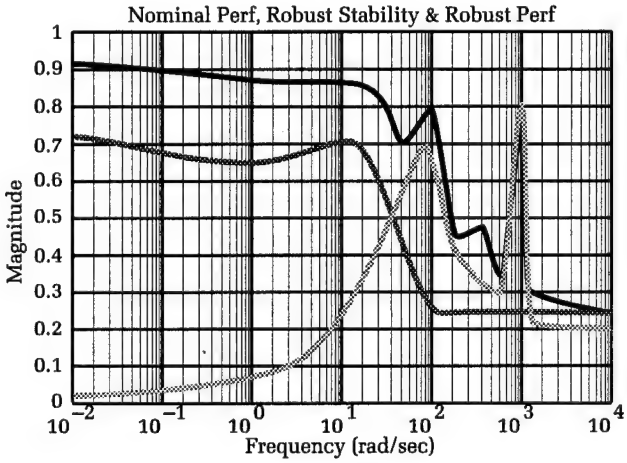


Figure 19. *Nominal performance, robust stability, and robust performance: reduced μ-design.*

Conclusions

The μ-synthesis robust control approach was applied to the design of a pitch plane autopilot for a high performance point defense missile considering flexible-body dynamics. The resulting control system was shown to provide robust stability and robust performance for the airframe, satisfying the time response and stability margin requirements, body bending mode attenuation requirements, and robustness criteria. The μ design was compared to a conventional autopilot design and was shown to provide superior time response characteristics.

The μ-synthesis design resulted in a 27-state controller which produced a time constant of 0.086 second with minimum overshoot. To simplify the design, the controller was then reduced to individual third-order controllers in each feedback path. This reduced-order design satisfied the performance specifications and robustness goals (including time response), while greatly simplifying the flight control system design, as well as the onboard computer implementation. This illustrates how a fairly complex controller can easily be reduced to manageable size with little or no performance penalty.

This study demonstrated several distinct benefits in using the μ-synthesis approach. First, and most importantly, the time constant of the flight control system can be reduced significantly. This allows the missile to react more quickly to low observable, fast maneuvering targets that do not allow much time for terminal homing. Second, robustness can be built into the controller to handle specific modeling and parameter uncertainties. Third, the μ controller serves as a benchmark by which conventionally designed autopilots may be checked, both for robust stability and robust performance. Finally, the μ-synthesis technique is a systematic approach to the design of robust flight controllers (including order reduction) that can directly incorporate performance requirements that are flowed down from higher level system trade studies. Thus, μ-synthesis provides a valuable tool in the design of flight control systems for ship air-defense missile systems.

Acknowledgments

The authors wish to thank Dr. H. L. Stalford for his contributions. This work was supported by the NSWCDD Independent Research program.

References

1. Nesline, F. William and Zarchan, P., "Robust Instrumentation Configurations for Homing Missile Flight Control," *Proceedings of the 1980 AIAA Guidance and Control Conference*, Danvers, MA, 1980.
2. Nesline, F. William and Nesline, Mark L., "How Autopilot Requirements Constrain the Aerodynamic Design of Homing Missiles," *Proceedings of the 1984 American Control Conference*, San Diego, CA, Jun 1984.
3. Nesline, F. William and Nesline, Mark L., "Homing Missile Autopilot Response Sensitivity to Stability Derivative Variations," *Proceedings of the 1984 IEEE Decision and Control Conference*, Las Vegas, NV, Dec 1984.
4. Cronvich, L. L., "Aerodynamic Considerations for Autopilot Design," Chapter 1 of *Tactical Missile Aerodynamics*, edited by M. J. Hemsch and J. N. Nielsen, AIAA Progress in Astronautics and Aeronautics Series, Vol. 104, 1986.
5. Ohlmeyer, E. J., *Analysis of Third-Order Autopilots for Missile Flight Control*, NAVSWC TR 90-335, NAVSWC, Dahlgren, VA, Dec 1990.
6. Reichert, R. T., "Robust Autopilot Design Using μ-Synthesis," *Proceedings of the 1990 American Control Conference*, San Diego, CA, May 1990.
7. Wise, K. A., Mears, B. C. and Poolla, K., "Missile Autopilot Design Using H_∞ Optimal Control with μ-Synthesis," *Proceedings of the 1990 American Control Conference*, San Diego, CA, May 1990.
8. Bibel, J. E. and Stalford, H. L., "Preliminary Missile Autopilot Design Using μ-Synthesis," *Proceedings of the Innovative Anti-Air Weapons Systems Conference*, JHU/APL, Jun 1990.
9. Ruth, M. J., "Robust Control of a Bank-to-Turn Missile," *Proceedings of the 1990 AIAA Guidance, Navigation and Control Conference*, Portland, OR, Aug 1990.
10. Bibel, J. and Stalford, H., "μ-Synthesis Autopilot Design for a Flexible Missile," *AIAA 91-0586, 29th Aerospace Sciences Meeting*, Reno, NV, Jan 1991.
11. Enns, D. F., "Rocket Stabilization as a Structured Singular Value Synthesis Design Example," *IEEE Control Systems Magazine*, Jun 1991.
12. Bibel, J. E. and Stalford, H. L., "An Improved Gain-Stabilized Mu-Controller for a Flexible Missile," AIAA-92-0206, *30th Aerospace Sciences Meeting and Exhibit*, Reno, NV, Jan 1992.
13. Dorato, P., A Historical Review of Robust Control," *IEEE Control Systems Magazine*, April 1987.
14. Ohlmeyer, E. J., *Robust Control Theory: Current Status and Future Trends*, NAVSWC MP 90-385, NAVSWC, Dahlgren, VA, Jun 1990.
15. Doyle, J. C., Lecture Notes in Advances in Multivariable Control, ONR/Honeywell Workshop, Minneapolis, MN, 1984.

16. Stein, G., "Beyond Singular Values and Loop Shapes," *AIAA Journal of Guidance, Control and Dynamics*, Vol. 14, No. 1, Jan-Feb 1991.
17. Doyle, J. C., "A Review of μ for Case Studies in Robust Control," *Selected Papers from the 10th Triennial World Congress of the IFAC*, Ed. R. Isermann, Pergamon, Oxford, UK, 1988.
18. Packard, A. and Doyle, J. C., *Robust Control of Multivariable and Large Scale Systems*, AFOSR-TR-88-0450, Honeywell Systems and Research Center, Minneapolis, MN, Mar 1988.
19. Doyle, J., "Analysis of Feedback Systems with Structured Uncertainties," *IEEE Proceedings*, Vol. 129, Part D, No. 6, Nov 1982.
20. Doyle, J. C., Wall, J. E. and Stein, G., "Performance and Robustness Analysis for Structured Uncertainty," *IEEE CDC Conference Proceedings*, 1982.
21. Doyle, J. C., "Structured Uncertainty in Control Systems Design," *IEEE CDC Conference Proceedings*, 1985.
22. Glover, K. and Doyle, J. C., "State-Space Formulae for all Stabilizing Controllers that Satisfy an H_∞ Norm Bound and Relations to Risk Sensitivity," *Systems and Control Letters*, Vol. 11, 1988.
23. Doyle, J. C., Glover, K., Khargonekar, P. P. and Francis, B. A., "State-Space Solutions to Standard H2 and H∞ Control Problems," *IEEE Trans. Auto. Control*, Vol. AC-34, No. 8, 1989.
24. Zames, G., "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," *IEEE Transactions of Automatic Control*, Vol. AC-26, 1981.
25. Bibel, J. E. and Malyevac, D. S., *Guidelines for the Selection of Weighting Functions for H-Infinity Control*, NSWCDD/MP-92/43, NSWCDD, Dahlgren, VA, Jan 1992.

The Authors

JOHN E. BIBEL received a B.S. degree in Aerospace Engineering from the Pennsylvania State University in May 1984. Since then, he has been employed at NSWCDD, where he has been involved with simulation; performance analysis; and guidance, navigation, control, and estimation technologies for tactical missile systems. Mr. Bibel is currently completing requirements toward an M.S. degree in Aerospace Engineering from Virginia Polytechnic Institute and State University and is a member of the American Institute of Aeronautics and Astronautics.

D. STEPHEN MALYEVAC received B.S. and M.S. degrees in Mechanical Engineering from Virginia Polytechnic Institute and State University in July 1986 and May 1988, respectively. He joined the Division in 1988 and works in the Aeromechanics Branch in the areas of guidance, control, simulation, and performance analysis for tactical missiles. He is a member of the American Institute of Aeronautics and Astronautics.

ERNEST J. OHLMEYER was born and raised in New Orleans, LA. He joined the Division in 1968 after receiving a B.S. degree in Physics from Loyola University. His early career focused on the free-flight dynamics of projectiles and design and development of the Navy 5-inch and 8-inch Guided Projectiles. In 1982, he obtained an M.S. degree in engineering from the Naval Postgraduate School as part of the Full-Time Advanced Study Program. His subsequent activities have emphasized analysis and performance evaluation of tactical missile systems, including Tomahawk, Wide Area Defense missile, Standard Missile, Surface Launched Weapon Technology, Advanced Self Defense Missile and Tactical Ballistic Missile Defense programs. Mr. Ohlmeyer has twice been a Principal Investigator in the NSWCDD Independent Research Program. Since 1985, he has been Group Leader for flight dynamics, guidance, and control in the Aeromechanics Branch (G23). His current interests include: simulation and modeling, robust and nonlinear control, system design optimization, and guidance algorithms for tactical missiles. Mr. Ohlmeyer is a member of AIAA, IEEE, and Sigma Xi.

Water Barrier Ship Self-Defense Concept

Charles E. Higdon

The Naval Surface Warfare Center Dahlgren Division (NSWCDD) is developing technology for a ship self-defense concept that has the potential to be effective in defending Navy platforms against high-speed, low-flying antiship missiles (ASMs). The concept uses a wall of water to provide a low-cost, universal, self-defense system for ships. This wall of water, or water barrier, is formed from the shallow detonations of underwater explosives in order to protect the ship from attacking ASMs. This concept can be employed to stop or slow debris and warhead fragments from missiles killed at short range to preclude significant damage to the defending ship. Furthermore, the barrier would defeat the fuzing and structure of ASMs that have penetrated the inner self-defense layer. Close-in employment of the barrier concept would increase the engagement space of self-defense weapons and help reduce detection range requirements. This article discusses the critical issues associated with the barrier concept and the aspects of the particular technologies that demonstrate its viability.

Introduction

Antiship missiles have become an enormous threat to ships over the past decade, as demonstrated by the 1982 Falklands Conflict, the 1987 attack on *USS Stark* by Iraq, and the proliferation of weapon technologies with the breakup of the Soviet Union.[1] With the expansion of antiship weaponry to many third-world nations and the increase in the Navy's requirement to operate in coastal waters, Navy platforms must be capable of defending themselves against a wide variety of ASMs that are launched from land, surface ships, submarines, and aircraft. As illustrated in Figure 1, many of these missiles use low-altitude, run-in trajectories to delay shipboard detection as long as possible, so as to decrease the battle space that is available to the defending ship. The detection range of low-altitude threats can be further decreased through signature reduction and the combined effects of multipath and clutter. These threats will be highly maneuverable, with subsonic to supersonic run-in speeds. The timing of the ASM launches may be coordinated to saturate the ship's antiair warfare (AAW) defenses and may be heavily supported by advanced jamming techniques. Thus, the sea-skimming ASM represents the most severe challenge to shipboard defense today.

Beginning in FY 93, the Surface Launched Weaponry (SLW) Technology Program, sponsored by the Office of Naval Research (ONR), funded the Self-Defense Weapon Technologies Project at NSWCDD to define and develop concepts and technologies that are required in providing a robust, all-weather, low-cost, self-defense system for a variety of ships. A wide-ranging review of technologies with the potential to support these goals resulted in identifying

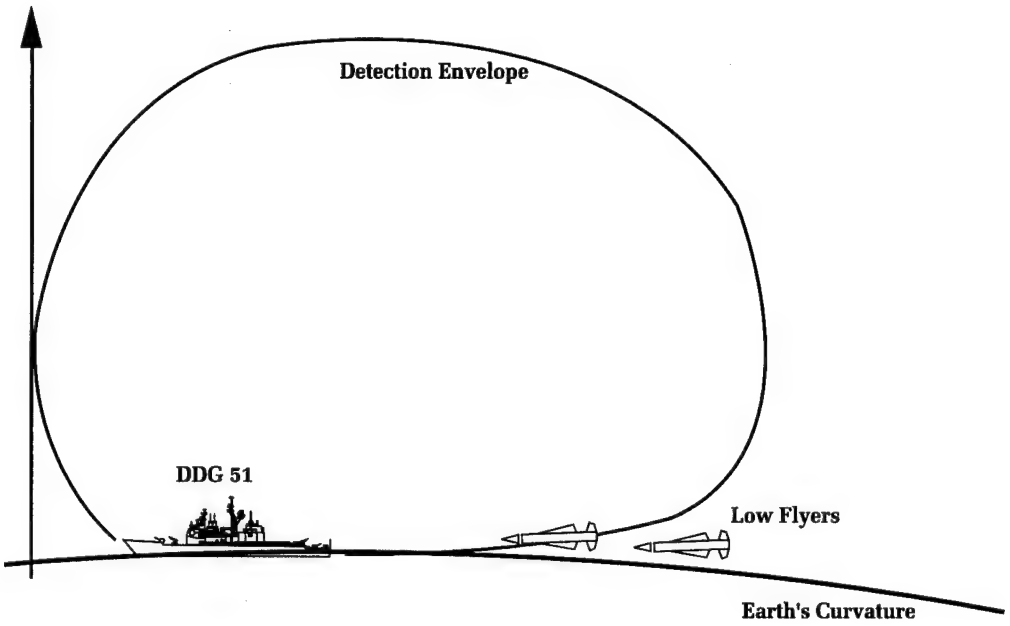


Figure 1. *The ASM threat attempts to delay shipboard detection through maneuvering, low-altitude trajectories. The detection range of low-altitude threats can be further decreased through signature reduction and the combined effects of multipath and clutter. With the increase in worldwide cruise-missile proliferation and technology transfer, these low-altitude threats can be launched from surface ships, land, aircraft, or submarines.*

several technologies that are currently being pursued by the SLW Technology Program. The objective of the water barrier ship self-defense concept is to improve AAW self-defense effectiveness against the supersonic, highly maneuverable, sea-skimmer threat at an affordable cost.

Barrier Concept Description

A concept having the potential to exploit seawater as a defense against low-flying ASMs uses a wall of water generated from underwater explosives to provide a low-cost, universal self-defense system for Navy ships.[2] The potential for using seawater to defend surface ships from air attack has been recognized several times in past AAW conflicts during and since World War II. The high density of water relative to air (1000 times) makes water an attractive medium if it can be placed in front of low-flying aircraft. The problem with previous approaches has been the inability to accurately place water in the path of these low-flying aircraft. The water barrier concept seeks to overcome this problem by creating a wall of water near the defending ship that blocks the path of the missile at the only point the missile path is known with certainty. The ship must be far enough from the wall of water, such that the ship is safe both from the shock effects of creating the barrier and from any penetrating debris or fragments, which should splash harmlessly into the sea.

This wall of water, or water barrier, is formed from the shallow detonations of underwater explosives to protect the ship from attacking, low-flying ASMs. The water barrier would provide protection to a ship under attack from the fragments of close-in intercepts by missiles or guns. In this case, the barrier would slow or stop the intercept debris and spare the ship from significant damage. Furthermore, the barrier concept has the potential to be effective against low-flying ASMs that penetrate the inner self-defense layer. The impact between a high-speed ASM and a dense water plume created by an underwater explosive charge near the water surface has the potential capability to:

- break up the ASM airframe
- significantly alter the ASM trajectory through structural damage to the radome or control surfaces
- detonate a fuzed warhead
- cause flameout of an air-breathing propulsion system

Antimissile missiles are typically very expensive, because they usually require a seeker to hit a target. However, seekers are vulnerable to countermeasures, multipath, clutter, glint, and noise. In fact, huge

technology investments are made to reduce seeker costs and increase their effectiveness. The barrier concept seeks to reduce cost and increase defensive robustness by eliminating the seeker. Placing a missile round at the predicted intercept point in the flight path of an ASM requires some form of minimal guidance. Without a seeker, multipath and clutter phenomena, and advanced jamming techniques would have no effect on the guidance of the barrier round. A round with no seeker and minimal guidance, using available shipboard detection and tracking resources, would provide a low-cost, self-defense system, when compared with the traditional self-defense missiles launched against low-flying ASMs.

In the barrier-defense operational mode, the Close-In Weapon System (CIWS) of Figure 2 will provide the normal detection and engagement of the sea skimmer, as depicted in Figure 3. Normally, CIWS will achieve multiple hits on the ASM, inducing the threat warhead or rocket motor to react violently. These hits will cause the incoming ASM (at speeds of Mach 1 or higher) to break into warhead fragments and missile debris. If the ASM breakup occurs close to the ship, this debris has the potential to inflict great damage to lightly armored modern ships. A deployed barrier or a wall of water would defeat the debris from close-in intercepts to protect the ship from greater damage. Furthermore, if the ASM is not destroyed at a critical range, the water barrier is deployed to form a wall of water to protect the ship from the attacking low flyer. Based on inputs from CIWS and existing shipboard detection and tracking systems, sufficient rounds are launched in a row perpendicular to the threat attack direction to mask the ship from the sea skimmer. These rounds (with underwater explosive charges) fly to a safe standoff distance of 300 to 450 feet from the ship. The rounds are detonated nearly simultaneously at a planned shallow depth to form a continuous wall of water in the path of the ASM. The water will slow the fragments and missile debris to harmless velocities, which allows CIWS, short-range missiles, and even a high-energy laser to destroy the ASM closer to the ship. If the ASM penetrates the inner

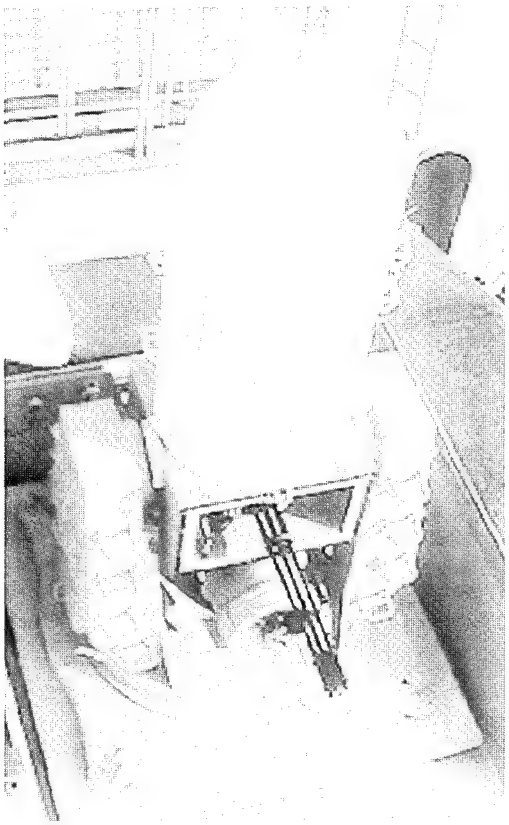
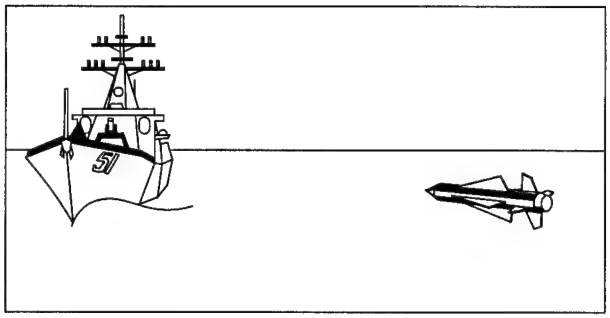
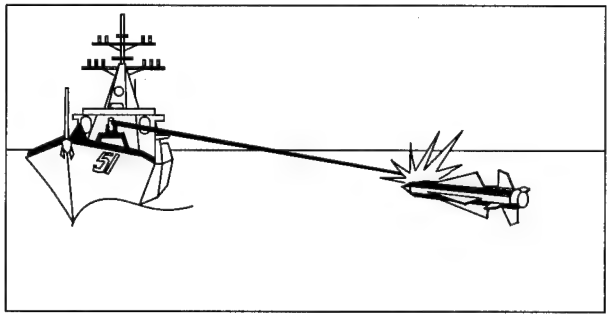


Figure 2. *The Phalanx CIWS is an M61A1 six-barreled Gatling gun designed to fire 4500 rounds per minute of saboted tungsten projectiles against incoming missiles. This system has self-contained search radar and tracking radar for automatic engagement of the target.*

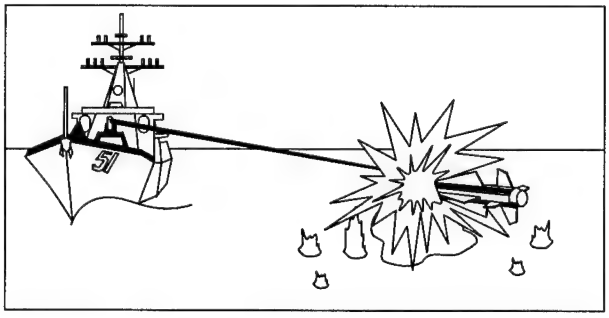
- Sea-skimming ASM attack detected
- Phalanx (or other short-range AAW system) engages normally



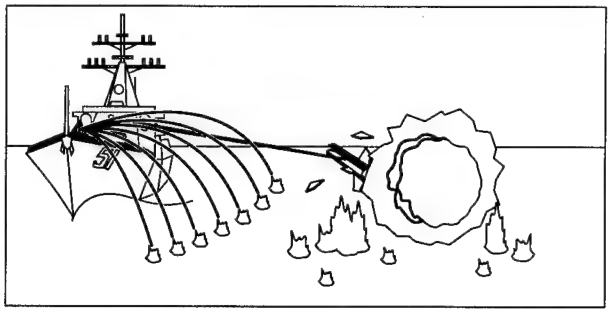
- Phalanx achieves multiple hits on the incoming ASM
- Barrier launcher trained to attack direction
- Rounds initialized (dive range, burst depth) based on known target characteristics



- Threat warhead (or rocket motor) reacts violently
- ASM breaks into warhead fragments and missile debris (which can cause critical damage)
- If ASM has not been destroyed by a critical range, barrier engagement is initiated
- CIWS engagement continues to minimum range



- Barrier system launches enough barrier rounds to mask ship from ASM attack direction
- Rounds fly to line perpendicular to attack direction at safe standoff from ship



- Barrier rounds detonate at planned depth
- Seawater thrown into path of ASM
- Water slows fragments to harmless velocities, allows CIWS to destroy ASM much closer to ship
- Additional ASMs trailing close behind (not engageable by CIWS) are destroyed by water barrier

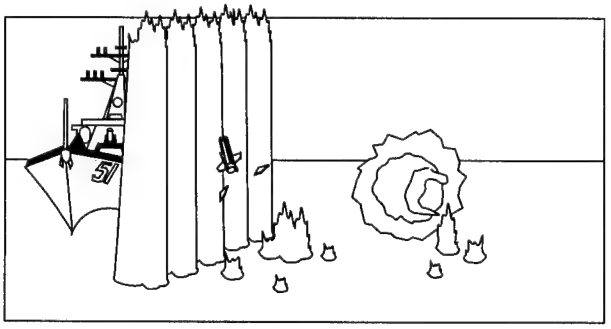


Figure 3. *Water barrier engagement sequence.*

self-defense layer, then the barrier will defeat the fuzing or structure of the ASM.

Underwater Explosion Characteristics

During the past 50 years, much research has been focused on the phenomenon of plumes generated from underwater explosives. A review of this research provides observations and empirical formulas (for single large charges) that describe the general shape of the resulting plume.[3-5] Despite this abundance of research, little is known quantitatively about the overall plume structure. The lack of detailed plume knowledge is due primarily to the opaqueness of the spray caused by the initial "spalling" of the surface from the explosive shock and the breakup of the ejected plumes. A hydrodynamic model has been developed to increase the understanding and knowledge of the plume structure and to aid in the development of the barrier concept.[6-7]

The most suitable index for describing an underwater explosion is the calculated maximum radius of the spherical explosion bubble, A_1. The equation for the radius of the fully expanded bubble is given by:

$$A_1 = J\frac{W^{\frac{1}{3}}}{Z^{\frac{1}{3}}} \qquad (1)$$

where:

A_1 = maximum radius of explosion bubble, ft
J = bubble radius coefficient
W = charge weight of explosive, lb
Z = hydrostatic pressure at depth of explosion, ft, usually expressed as $(d + 33)$, with d equal to the explosion depth in feet

Figure 4 depicts the complex nature of the plume phenomena from a shallow underwater explosion. This figure is a two-dimensional hydrodynamic time-sequence representation of a single 25-pound charge of PBXN-103 detonated at a depth of 10 feet. A shallow underwater explosion is defined for a scaled depth (d/A_1) that is less than 1.0 (the scaled depth is 0.6 in this example). The plume densities within the plume structure are represented by the shaded scale on the right-hand side in the figure, where full density of water is represented by the dark shading. When an explosion is shallow ($d < A_1$), the layer of water above the explosive charge is pushed up at a high velocity just as the first bubble begins to expand after detonation. As the bubble begins to collapse, the water flowing upward converges to form a central, vertical water jet within the plume and bubble. This central jet within the plume rises to a relatively great height. As the second bubble expands and collapses, dense secondary, or radial, plumes are formed around the existing central, vertical water jet. The plume remains extremely dense in the lower regions for over 3 seconds. The dense lower regions of the plume (< 30 feet) will be exploited as a self-defense weapon against low-flying ASMs.

Fragment/Missile Encounter

In order to exploit the plume as a self-defense weapon, the expected effects of water on fragments and missiles must be examined. Figure 5 shows the expected velocity loss for chunky steel fragments as a function of distance traveled through 100 percent water and fragment weight. These curves are developed from the exponential drag model used by the Joint Technical Coordinating Group for Munitions Effectiveness to predict fragment velocity loss in fluids. Considering a large warhead fragment size of 3000 grains, a 70-percent velocity loss of the strike velocity requires only 1.5 feet of full-density water. Likewise, a 70-percent velocity loss of the strike velocity for a large 1-pound fragment requires only 2 feet of water. Thus 1.5 to 2 feet of full-density water can drastically reduce fragment velocities generated from close-in threat intercepts.

The loads on a missile are proportional to the dynamic pressure. Small increases in densities will thus amplify the loads on the missile structure, with the potential to cause failures. When a missile encounters water, the dynamic pressure on that missile will increase significantly. Dynamic pressure is defined in the following equation as:

$$q = \frac{1}{2}\rho V^2 \qquad (2)$$

where:

q = dynamic pressure of the medium
ρ = density of the medium
V = velocity of the object

Figure 6 shows the instantaneous amplification of the dynamic pressure in air as a function of the amount of water in a given

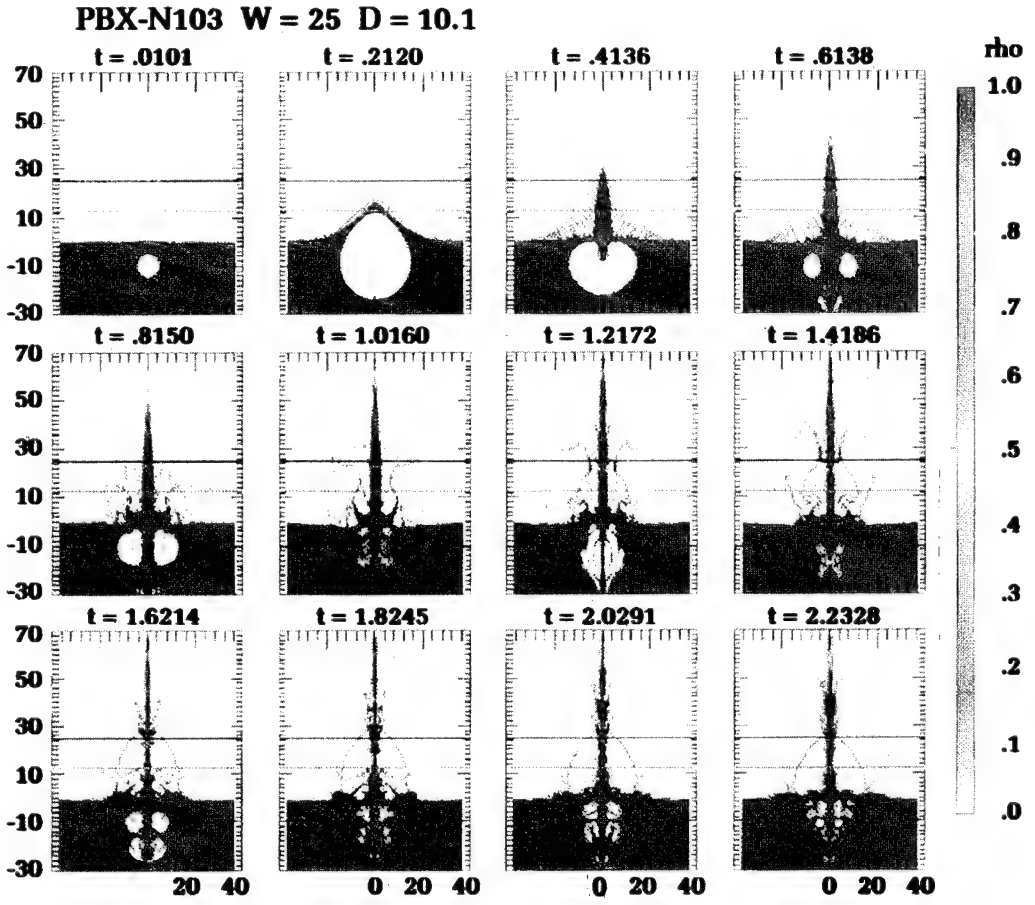


Figure 4. *Hydrodynamic calculation of water plume from detonation of a 25-pound charge of PBXN-103 explosive at a depth of 10 feet.*

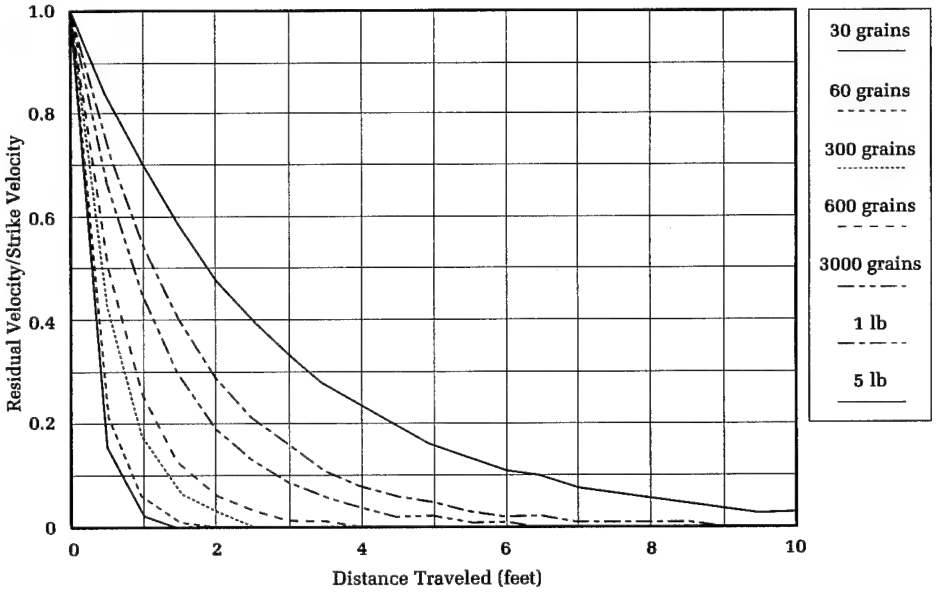


Figure 5. *Expected velocity loss, chunky steel fragments, 100 percent water.*

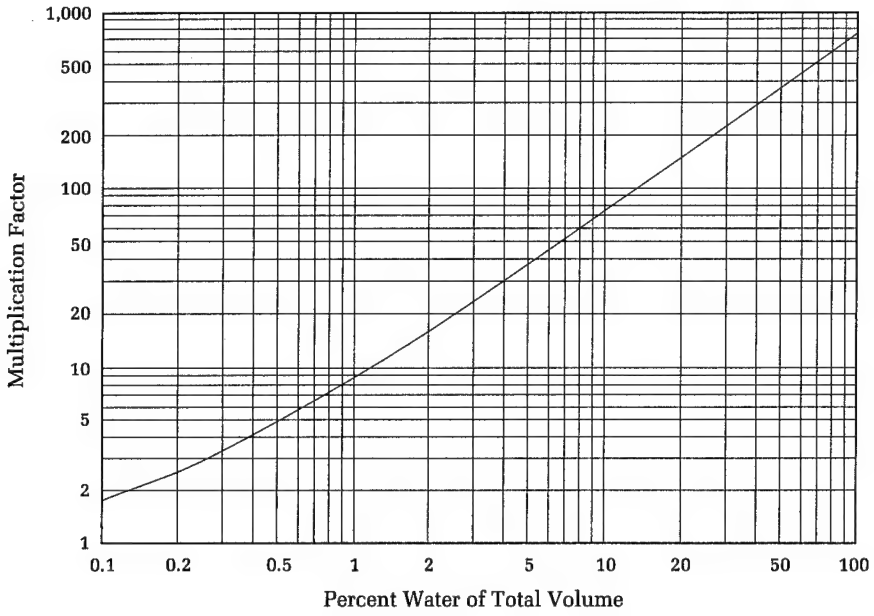


Figure 6. *Dynamic pressure amplification.*

volume, where 100 percent represents full-density water, and zero percent represents air only. An examination of a hypothetical missile with a hypothetical 15-G maneuvering capability shows that a small percentage of water in a given volume can dramatically increase the loads on that missile. The maximum design load for a 15-G missile is approximately 1.25 to 1.5 times 15 Gs, for a maximum load capability of 22.5 Gs. In the nonmaneuvering case, or 1-G level, 3 percent of water in air will amplify the load on the missile to approximately 25 Gs, which is beyond the design limit of the missile. For the 15-G maneuvering case, just over 0.1 percent of water in air will double the load on the missile to 30 Gs. Therefore, a small percentage of water in air (0.1% to 3%) can readily increase the loads on a missile that are beyond its design limits. These increased loads can lead to the total breakup and destruction of the missile and to flight-path direction changes, diverting the missile from its path toward the ship.

When a missile impacts the water barrier, significantly large longitudinal decelerations of 600 to 6000 Gs are expected, depending on threat velocities and mass.[2] Pressure loadings on the radome and the leading edges of the airframe are on the order of 10^6 to 10^7 psf, with large impact shocks transmitted into the airframe. Therefore, the expected result of a threat missile encounter with a water barrier close to the ship is warhead detonation by fuze activation. Because the water barrier has dimensions similar to the ship, the expected detonation would occur inside the barrier, where the fragments and debris are slowed or stopped to drop harmlessly into the sea. Predictions based on finite element analysis appear to support predictions of structural failure and high longitudinal loads when a missile encounters the water barrier.[8]

Barrier Field Tests

Two series of field tests were conducted in FY 93. The purpose of these tests was to validate the hydrodynamic model, to develop plume measurement techniques, to characterize the plume structure, and to form a water barrier from scale charges of HE. The capability to predict and quantify the plume phenomena from shallow detonations of underwater explosives is crucial in exploiting the plume as a self-defense weapon. Progress in this ability requires the validation and refinement of the hydrodynamic model to accurately predict plume performance. To aid the validation process requires developing plume density measurement techniques to experimentally evaluate the internal plume structure. Conductivity probes and microwaves are two complementary plume-measurement techniques that were designed to experimentally ascertain the equivalent length of 100 percent water within a plume as a function of time.[9,10] Finally, the plume structure can be characterized using both the hydrodynamic model and experimental data from the field tests.

Small-Scale, Single-Charge Tests

The first series of FY 93 field tests consisted of 1-pound, small-scale charges of C-4 and pentolite. These HE charges were tested at various depths at the NSWC Carderock Division Test Pond in Bethesda, Maryland. Figure 7 shows the photographs of the plume and bubble from Shot 9 during the initial stages of plume development. The explosive is C-4, with a charge weight of 1.25 pounds detonated at a depth of 2.5 feet. The time in the formation of the plume and bubble is at 100 milliseconds from detonation, which is at approximately the maximum bubble radius. Figure 8 compares the hydrodynamic model with photographic data of both the plume and bubble from Shot 9 at 100 milliseconds from detonation. The solid line in the figure is the 96-percent density contour of both the plume and bubble as predicted from the hydrodynamic model. This contour shows the central jet formation within the plume and bubble. The dashed line is the 0.1-percent density contour as generated from the model. This density contour is an approximation of what is optically seen. The n symbol represents the plume and bubble profiles as obtained from the photographic observations. This graph clearly shows that the model computations for the 0.1-percent density contour match up quite well with the observed photographic data.

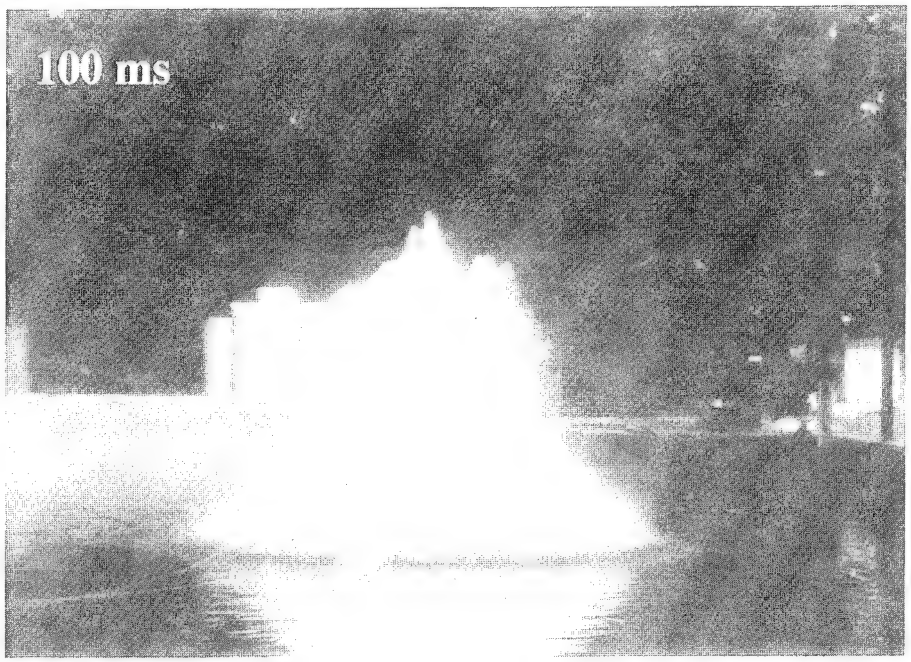
Intermediate-Scale, Single- and Multiple-Charge Tests

The second series of FY 93 field tests consisted of single and multiple 10-pound scale charges of C-4 that were tested in the quarries at Hi-Test Laboratories, Arvonia, Virginia. Multiple charges were tested at various depths and linear spacings to form a water barrier. During these barrier experimental tests, density measurements of the resulting plumes were made using conductivity probes and microwaves. Figure 9 compares, to scale, a single charge of C-4 with a barrier created from five charges of C-4. The single charge was detonated at a depth of 10 feet. The barrier charges are spaced at 8 feet apart and at a depth of 8 feet. The photographs of each shot were taken just before maximum plume height. The plume from a single charge shows a height of approximately 27 feet, with a plume-base diameter of approximately 37 feet. In comparison, the barrier central plume has a height of 70 feet, which grows to a maximum of nearly 100 feet. The secondary, or radial, plume heights, which are located to the front and rear of the central barrier plume, obtain heights of almost 50 feet. The length of the barrier is over 40 feet and maintains this length to almost the secondary plume heights. The thickness (length perpendicular to the charge line) of the barrier plume, which includes the central plume and the front and rear radial plumes, is over 60 feet. Thus, the multiple charges greatly amplify the single-plume dimensions to form a water barrier.

Using a barrier configuration of five multiple charges of C-4 with a charge depth of 8 feet and a spacing of 8 feet between the charges, Figure 10 compares the integrated densities of the microwave data (solid line) with the computations from the hydrodynamic model (dashed line) at a height of 25 feet above the water surface. The integrated density (expressed as the equivalent length of 100% water) is plotted as a function of time. The microwave data is truncated at about 1.5 feet, because values above this point were beyond the dynamic range of the microwave equipment. At a minimum, the microwave data shows that the barrier plume has an equivalent water length of ≥1.5 feet at a height of 25 feet above the water surface. The 1.5 feet or more of equivalent water length lasts for over 2 seconds, or nearly 0.5 second longer than the model prediction. This equivalent water length (≥1.5 feet) is predicted to represent a 70 percent loss of strike velocity for 3000-grain fragments (see Figure 5). In comparing the experimental data with the hydrodynamic model, the rise and fall times for the barrier plume seem to be basically the same. The hydrodynamic model indicates that the barrier has an equivalent water length of 2 feet, with a duration of 1 second. This equivalent water length is predicted to represent a 70-percent loss of strike velocity for 1-pound fragments.

These results are for equivalent water lengths at a 25-foot height above the water surface and extend along the length of the barrier. For heights less than 25 feet, the equivalent water lengths should be longer, with longer durations. In addition, the predicted and experimental data of the barrier plume at a 25-foot height seem to indicate that 10-pound charges generate sufficient water to stop or slow large fragments.

The FY 93 field test results show that the hydrodynamic model appears to be consistent with the experimental data. The multiple charges of C-4 amplify single-charge plume dimensions to form a water barrier. Even a barrier of 10-pound scale charges appears to have potential tactical value as a self-defense weapon in stopping or slowing



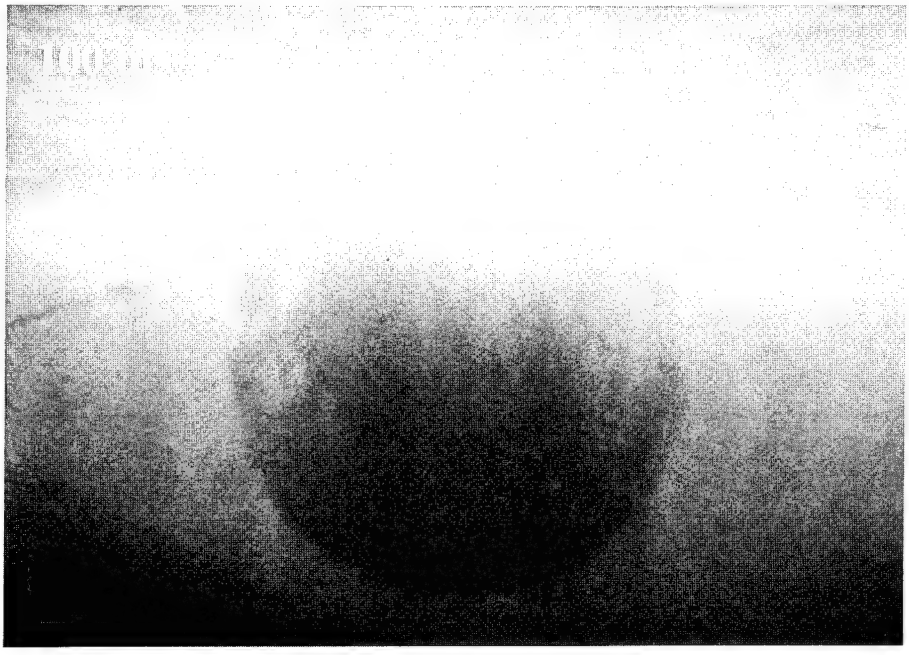


Figure 7. *Photos of bubble and plume of Shot 9 at Carderock.*

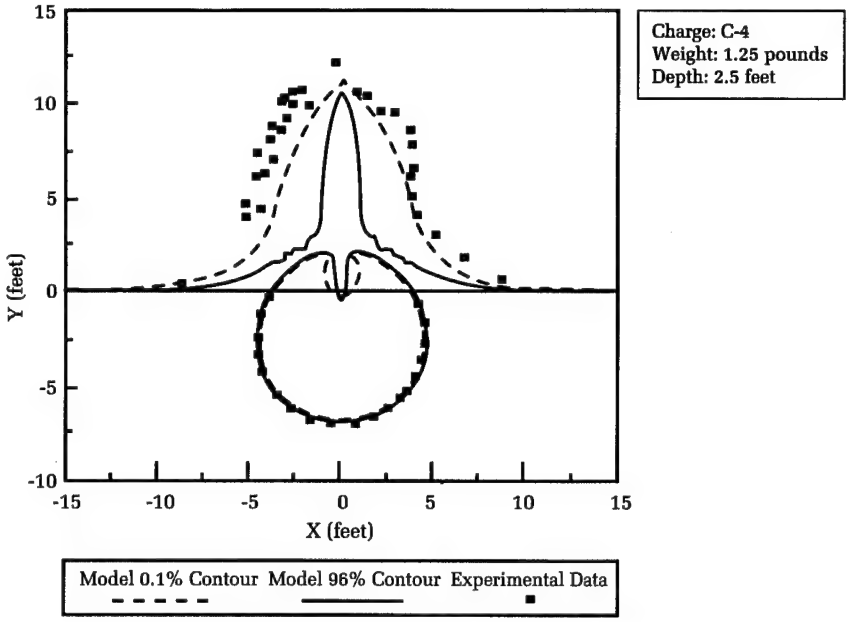


Figure 8. *Density contour (100 ms), Shot 9—Carderock Test Pond, May 1993.*

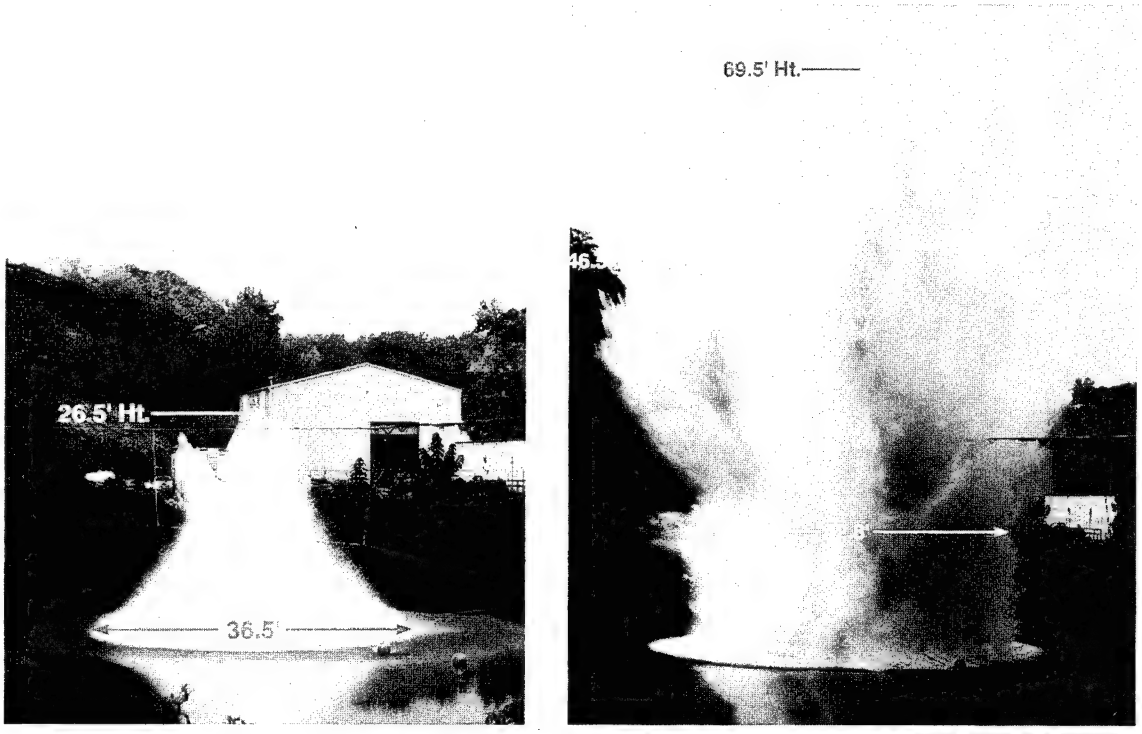


Figure 9. *Single vs. barrier plume, Arvonia, VA, August 1993.*

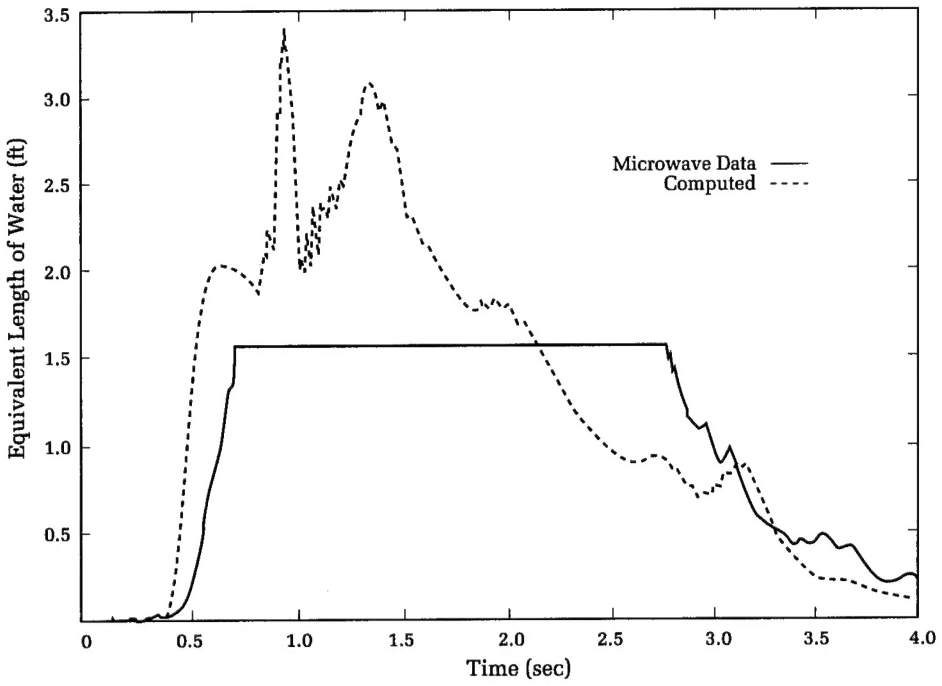


Figure 10. *Arvonia Shot 3—model comparison with microwave data at height of 25 feet.*

large fragments. The initial water barrier concept anticipated the need for 100-pound equivalent TNT charges to create a barrier. Based on these results, the barrier might be effective with explosive charges far smaller than 100 pounds, due to the multiple-charge amplification effect. As a consequence, using a smaller charge for a barrier leads to a reduced charge weight and size, a smaller round, easier insensitive munition requirements, and less shock received at the ship.

Ability of Barrier to Stop Fragments and Projectiles

A subsequent series of shallow underwater explosives trials were successfully concluded during June/July 1994 at the Briar Point Test Pond at the Aberdeen Proving Ground, Aberdeen, Maryland. The purpose of these trials was to:

- determine the ability of a water-barrier plume to stop or slow high-velocity projectiles and fragments for ship self-defense,
- refine the hydrodynamic model predictions,
- acquire inputs for alternative explosive types, and
- investigate larger scale explosive effects on barrier-plume predictions.

These tests involve the nearly simultaneous detonation of five 25-pound charges of PBXN-103 and PLGUW-18F at depths and spacings defined through hydrodynamic analysis to maximize the quantity of water thrown into the air. The initial phase of these tests characterized the plume formation and dimensions with high-speed photography and directly measured the plume structure using conductivity probes and microwaves. Preliminary results indicate that the barrier formation is somewhat robust to variation in charge depth, spacing, and even the failure of a single charge.

Figure 11 shows the development of the barrier plume from detonation to approximately its maximum height. Figure 11(a) shows the barrier plume just after the nearly simultaneous detonation of five 25-pound charges of PBXN-103. The charge depth is 10 feet, with a separation of 10 feet between the charges. The four protrusions from the barrier plume are the result of the shock interaction between the charges. These protrusions are located approximately midway between the charge positions. The plume height at this time is approximately 22 feet. Density measurements of the plume are made using the conductivity probes supported by the steel cable in the figure and the microwave antennas mounted on the tower in the background.

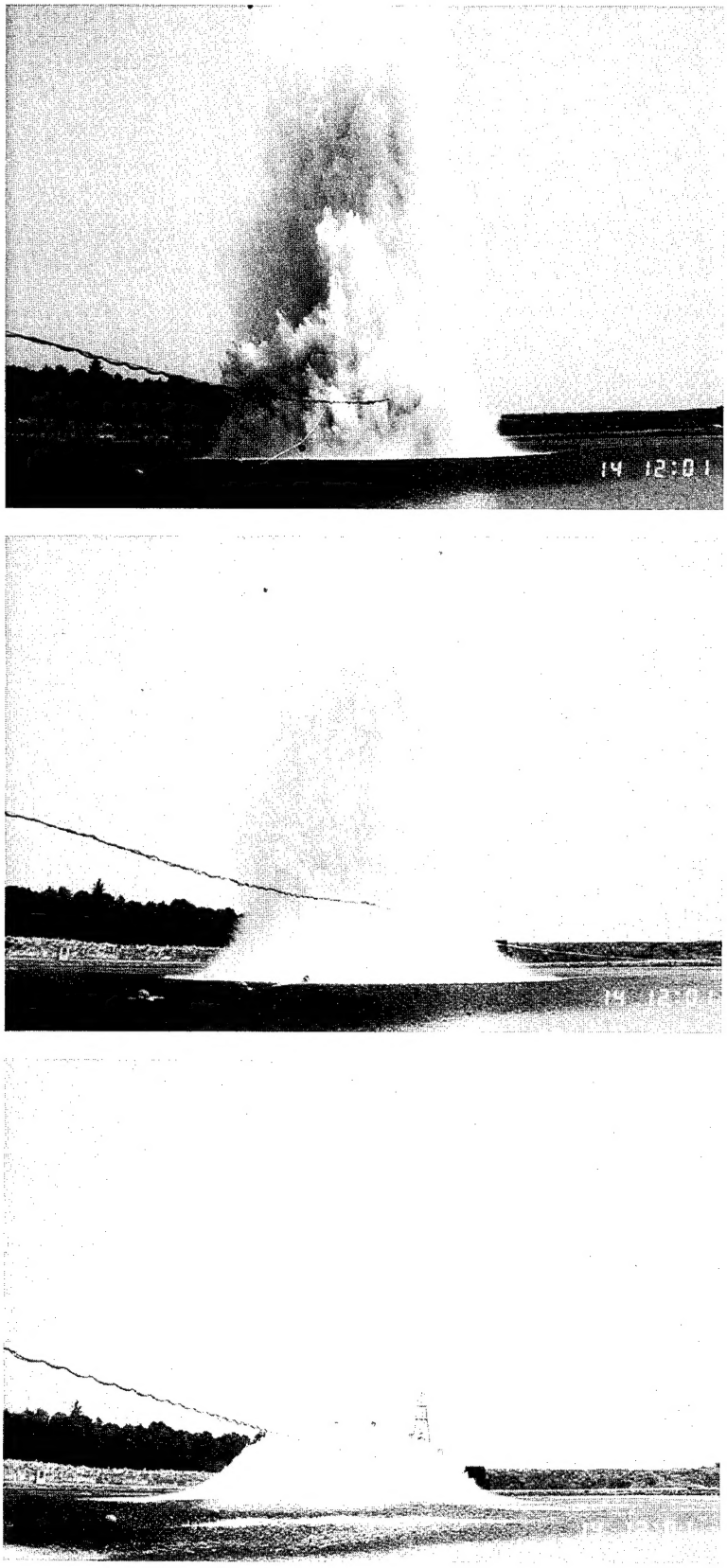


Figure 11. *Barrier development from near-simultaneous detonation of five 25-pound charges of PBXN-103 explosive spaced 10 feet apart and 10 feet deep.*

When a row of multiple charges are detonated nearly simultaneously, a cylindrical bubble is formed. As the first cylindrical bubble begins to collapse, a dense central barrier plume is generated that rises to a great height. In Figure 11(b), the central plume has attained a height of nearly 130 feet, with a width of about 80 feet. As the second cylindrical bubble expands and collapses, dense secondary plumes are formed in front of and behind the central plume. In Figure 11(c), the central plume height is over 140 feet, with the width still about 80 feet. The secondary plumes are just under 75 feet.

The final phase of the test included a 155-mm gas gun firing a ballasted nylon projectile (simulating missile ballistic density) and multiple .50-caliber Mann guns firing fragment-simulating projectiles. When the test projectiles and fragments encountered the water barrier within 0.5 to 2.5 seconds of charge detonation, they did not penetrate the barrier. For encounters that occurred between 2.5 and 3.5 seconds after charge detonation, the fired projectiles and fragments that were able to penetrate the barrier showed greatly reduced velocities and fell harmlessly into the pond.[11]

Future Efforts

While the water barrier phenomenon appears to have the potential to be effective, further work is required to scale up the 155-mm tests to full missile size and to supersonic speeds. Predictions based on finite-element analysis appear to support predictions of structural failure and high longitudinal loads. Tests with real fuzes are required to validate fuze function predictions and to assess the fragment-stopping abilities of the barrier where warhead detonations take place inside the barrier plume. Finally, refinement of a system concept to exploit the phenomenon must be pursued.

Conclusions

The water barrier ship self-defense concept has the potential to be particularly effective in defending Navy platforms against high-speed, low-flying ASMs. This concept's advantage is the ability of a wall of water to stop or slow debris and warhead fragments from missiles killed at very short range to preclude significant damage to the defending ship. Furthermore, the barrier would defeat the fuzing and structure of ASMs that have penetrated the inner self-defense layer. The ability to predict and quantify plume phenomena is crucial in exploiting the barrier plume as a ship self-defense weapon. Consequently, two techniques have been developed to experimentally measure the internal structure of the barrier plume. These measurement techniques were used to refine and validate the hydrodynamic model to better predict barrier-plume performance. Comparison of the model and the experimental data collected in the field tests shows that the hydrodynamic model appears to be consistent with the experimental data. This comparison has advanced the knowledge and understanding of the plume structure generated by underwater explosives detonated at shallow depths.

Observations of the barrier field tests have shown that multiple explosive charges of 10 to 25 pounds can successfully form a wall of water. The multiple HE charges of underwater explosives greatly amplify single-charge plume dimensions. Furthermore, the underwater explosives of PBXN-103 and PLUGW-18F have successfully stopped or slowed multiple .50-caliber fragments and a 155-mm ballasted nylon projectile (simulating missile ballistic density). A water barrier using 25- to 30-pound charges of HE appears to be tactically useful as a ship self-defense weapon.

Acknowledgments

The author wishes to acknowledge the significant contributions of the following individuals in the development of the water barrier concept and the determination of its effectiveness. Mr. Thomas McCants of the Ship Self-Defense Department at NSWCDD provided major inputs in the development and systems engineering of the water barrier concept. Mr. Len Lipton of the Warhead Technology and Development Department at NSWC Indian Head Division developed the conductivity probe measurement technique and directed the complex and difficult series of barrier field tests to a successful conclusion. Dr. William Szymczak of the Systems Research and Technology Department at NSWCDD developed the hydrodynamics model that has led to an increase in the understanding of the complex phenomena of water plumes generated from shallow, underwater explosives. Dr. Joon Choe and Mr. Kevin Boulais of the Weapons Research and Technology Department at NSWCDD have developed a technique for the measurement of density in water plumes using microwaves.

References

1. Zaloga, S., "HARPOONSKI," *U.S. Naval Institute Proceedings*, Feb 1994, pp. 37-40.
2. Higdon, C. E., *Water Plume Concept for Ship Self Defense*, NAVSWC TR 91-710, Naval Surface Warfare Center, Silver Spring, MD, Oct 1991.
3. Cole, R. H., *Underwater Explosions*, Princeton University Press, Princeton, NJ, 1948.
4. Young, G. A., *Plume and Ejecta Hazards from Underwater Explosions*, NOLTR 73-111, Naval Ordnance Laboratory, White Oak, MD, Dec 1973.
5. Young, G. A., *Dispersions of Chemical Products of Underwater Explosions*, NSWC TR 82-404, Naval Surface Warfare Center, Silver Spring, MD, Dec 1984.
6. Szymczak, W. G. and Wardlaw, A. B., *Numerical Methods for Explosion Predictions*, NAVSWC TR 91-718, Naval Surface Warfare Center, Silver Spring, MD, Mar 1993.
7. Szymczak, W. G., *Computations and Experiments of Shallow Depth Explosion Plumes*, NSWCDD/TR-94/156, Naval Surface Warfare Center Dahlgren Division, White Oak, MD, Draft.
8. Miller, G. S., "Barrier Self-Defense: Missile Response Using a Simple Beam Model Presentation," NSWCDD, Dahlgren, VA, 8 Aug 1994.
9. Lipton, L. D., *Probe for Measurement of Water Mass of Plumes Produced by Underwater Detonations*, IHTR 1757, Naval Surface Warfare Center Indian Head Division, Indian Head, MD, Draft.
10. Choe, J., et al., "The Mass Measurement of Water Plumes Produced by Underwater Explosion," in *Technical Proceedings of the 1994 Test Technology Symposium (TTS-VII)*, Johns Hopkins University/ Applied Physics Laboratory, Laurel, MD, 29-31 Mar 1994, pp. 493-504.
11. Lipton, L. D., *Water Barrier Concept for Ship Self-Defense: Field Test Results*, IHTR 1758, Naval Surface Warfare Center Indian Head Division, Indian Head, MD, Draft.

The Author

CHARLES E. HIGDON, an electronics engineer in the Weapons Systems Department since 1967, holds a BSEE degree from the University of Evansville. His major work experience includes developing weapon system simulations that model the system characteristics, defining top-level system requirements for advanced weapon systems, and analyzing and developing advanced system concepts. Currently, he is working on the development of self-defense concepts and the characterization of concept parameters for ONR self-defense efforts.